



US 20070231303A1

(19) **United States**

(12) **Patent Application Publication**

**Roy et al.**

(10) **Pub. No.: US 2007/0231303 A1**

(43) **Pub. Date: Oct. 4, 2007**

(54) **METHODS OF GENERATING CHIMERIC ADENOVIRUSES AND USES FOR SUCH CHIMERIC ADENOVIRUSES**

(75) Inventors: **Soumitra Roy**, Wayne, PA (US); **James M. Wilson**, Gladwyne, PA (US)

Correspondence Address:  
**HOWSON AND HOWSON**  
**SUITE 210**  
**501 OFFICE CENTER DRIVE**  
**FT WASHINGTON, PA 19034 (US)**

(73) Assignee: **The Trustees of the University of Pennsylvania**, Philadelphia, PA (US)

(21) Appl. No.: **10/465,302**

(22) Filed: **Jun. 20, 2003**

**Publication Classification**

(51) **Int. Cl.**  
*A61K 48/00* (2006.01)  
*C07H 21/02* (2006.01)  
*C12N 15/861* (2006.01)  
*C12N 5/06* (2006.01)  
*C12N 7/00* (2006.01)  
*C07K 14/075* (2006.01)

(52) **U.S. Cl.** ..... **424/93.2**; 435/235.1; 435/456;  
435/362; 530/350; 536/23.72

(57) **ABSTRACT**  
A method for providing an adenovirus from a serotype which does not grow efficiently in a desired cell line with the ability to grow in that cell line is described. The method involves replacing the left and right termini of the adenovirus with the corresponding termini from an adenovirus which grow efficiently in the desired cell line. At a minimum, the left terminus spans the 5' inverted terminal repeat, the left terminus spans the E4 region and the 3' inverted terminal repeat. The resulting chimeric adenovirus contains the internal regions spanning the genes encoding the penton, hexon and fiber from the serotype which does not grow efficiently in the desired cell. Also provided are vectors constructed from novel simian adenovirus sequences and proteins, host cells containing same, and uses thereof.

## METHODS OF GENERATING CHIMERIC ADENOVIRUSES AND USES FOR SUCH CHIMERIC ADENOVIRUSES

### BACKGROUND OF THE INVENTION

[0001] The presence of humoral immunity (circulating antibodies) to adenovirus capsid proteins is a barrier to the use of adenovirus vectors for gene therapy. The prototype adenovirus vectors that have been developed for gene therapy are based on subgroup C adenoviruses such as that of serotype 5. The prevalence of neutralizing antibodies against subgroup C adenoviruses is generally high in human populations as a result of frequent exposure to these pathogens. This fact is likely to greatly limit the effectiveness of gene therapy vectors based on serotypes such as Ad5.

[0002] Analysis of the nature of the protective antibodies against adenoviruses has indicated that the most important target is the major capsid protein, hexon [Wolfhart (1988) *J. Virol.* 62, 2321; Gall et al. (1996) *J. Virol.* 70, 2116]. Several efforts have been made to engineer the hexon so as to evade the anti-hexon antibodies by making chimeric adenoviruses harboring hexons from other serotypes [Roy et al. (1998) *J. Virol.* 72, 6875; U.S. Pat. No. 5,922,315; Gall et al. (1998) *J. Virol.* 72, 10260; Youil et al. (2002) *Hum. Gene Ther.* 13, 311; Wu et al. (2002) *J. Virol.* 76, 12775]. However, this has been largely unsuccessful when exchanges among distant serotypes are attempted.

[0003] Alternatively, investigators have proposed using adenovirus vectors that rarely cause human infections or using adenoviruses from non-human sources. However, the lack of a practical manner in which to produce large numbers of such vectors has proved to be a hindrance to developing such vectors.

### SUMMARY OF THE INVENTION

[0004] The present invention provides a method of modifying adenoviruses having capsids, and particularly, including hexons, from serotypes which are not well adapted for growth in cells useful for adenoviral virion production. The method is useful for production of scalable amounts of adenoviruses. The modified, or chimeric, adenoviruses are useful for a variety of purposes which are described herein.

[0005] The invention further provides novel, isolated, adenovirus SA18 nucleic acid and amino acid sequences, vectors containing same, cell lines containing such SA18 sequences and/or vectors, and uses thereof.

[0006] Other aspects and advantages of the present invention will be readily apparent from the following detailed description of the invention.

### DETAILED DESCRIPTION OF THE INVENTION

[0007] The present invention provides chimeric adenoviruses composed of the left terminal end and right terminal end of an adenovirus which can be cultured in the selected host cell, and the internal regions encoding, at a minimum, the capsid proteins of another adenovirus serotype. This invention is particularly advantageous for generating adenoviruses having serotypes which are difficult to culture in a desired cell type. The invention thus permits generation of chimeric adenoviruses of varying serotypes.

[0008] In the embodiments illustrated herein, chimeric adenoviruses have been constructed where most structural proteins, and not merely the hexon or fiber, are derived from an adenovirus of an unrelated serotype, thereby preserving the majority of the protein-protein interactions that are involved in capsid assembly. Most of the early genes such as those encoded by the adenovirus E1 and E4 regions that are responsible for transcription regulation and regulation of the host cell cycle, are retained from a different serotype that is known to result in high titer virus generation in the commonly used cell types, such as HEK 293 which supplies the Ad5 E1 proteins in trans.

[0009] In another embodiment, the invention provides novel nucleic acid and amino acid sequences from Ad SA18, which was originally isolated from vervet monkey [ATCC VR-943]. The present invention further provides novel adenovirus vectors and packaging cell lines to produce those vectors for use in the in vitro production of recombinant proteins or fragments or other reagents. The invention further provides compositions for use in delivering a heterologous molecule for therapeutic or vaccine purposes. Such therapeutic or vaccine compositions contain the adenoviral vectors carrying an inserted heterologous molecule. In addition, novel sequences of the invention are useful in providing the essential helper functions required for production of recombinant adeno-associated viral (AAV) vectors. Thus, the invention provides helper constructs, methods and cell lines which use these sequences in such production methods.

[0010] The term “substantial homology” or “substantial similarity,” when referring to a nucleic acid or fragment thereof, indicates that, when optimally aligned with appropriate nucleotide insertions or deletions with another nucleic acid (or its complementary strand), there is nucleotide sequence identity in at least about 95 to 99% of the aligned sequences.

[0011] The term “substantial homology” or “substantial similarity,” when referring to amino acids or fragments thereof, indicates that, when optimally aligned with appropriate amino acid insertions or deletions with another amino acid (or its complementary strand), there is amino acid sequence identity in at least about 95 to 99% of the aligned sequences. Preferably, the homology is over full-length sequence, or a protein thereof, or a fragment thereof which is at least 8 amino acids, or more desirably, at least 15 amino acids in length. Examples of suitable fragments are described herein.

[0012] The term “percent sequence identity” or “identical” in the context of nucleic acid sequences refers to the residues in the two sequences that are the same when aligned for maximum correspondence. The length of sequence identity comparison may be over the full-length of the genome (e.g., about 36 kbp), the full-length of an open reading frame of a gene, protein, subunit, or enzyme [see, e.g., the tables providing the adenoviral coding regions], or a fragment of at least about 500 to 5000 nucleotides, is desired. However, identity among smaller fragments, e.g., of at least about nine nucleotides, usually at least about 20 to 24 nucleotides, at least about 28 to 32 nucleotides, at least about 36 or more nucleotides, may also be desired. Similarly, “percent sequence identity” may be readily determined for amino acid sequences, over the full-length of a protein, or a fragment thereof. Suitably, a fragment is at least about 8

amino acids in length, and may be up to about 700 amino acids. Examples of suitable fragments are described herein.

[0013] Identity is readily determined using such algorithms and computer programs as are defined herein at default settings. Preferably, such identity is over the full length of the protein, enzyme, subunit, or over a fragment of at least about 8 amino acids in length. However, identity may be based upon shorter regions, where suited to the use to which the identical gene product is being put.

[0014] As described herein, alignments are performed using any of a variety of publicly or commercially available Multiple Sequence Alignment Programs, such as "Clustal W", accessible through Web Servers on the internet. Alternatively, Vector NTI utilities are also used. There are also a number of algorithms known in the art that can be used to measure nucleotide sequence identity, including those contained in the programs described above. As another example, polynucleotide sequences can be compared using Fasta, a program in GCG Version 6.1. Fasta provides alignments and percent sequence identity of the regions of the best overlap between the query and search sequences. For instance, percent sequence identity between nucleic acid sequences can be determined using Fasta with its default parameters (a word size of 6 and the NOPAM factor for the scoring matrix) as provided in GCG Version 6.1, herein incorporated by reference. Similarly programs are available for performing amino acid alignments. Generally, these programs are used at default settings, although one of skill in the art can alter these settings as needed. Alternatively, one of skill in the art can utilize another algorithm or computer program that provides at least the level of identity or alignment as that provided by the referenced algorithms and programs.

[0015] As used throughout this specification and the claims, the term "comprise" and its variants including, "comprises", "comprising", among other variants, is inclusive of other components, elements, integers, steps and the like. The term "consists of" or "consisting of" are exclusive of other components, elements, integers, steps and the like.

[0016] Except where otherwise specified, the term "vector" includes any genetic element known in the art which will deliver a target molecule to a cell, including, naked DNA, a plasmid, phage, transposon, cosmids, episomes, viruses, etc.

[0017] By "minigene" is meant the combination of a selected heterologous gene and the other regulatory elements necessary to drive translation, transcription and/or expression of the gene product in a host cell.

[0018] As used herein, the term "transcomplement" refers to when a gene (gene product) of one adenovirus serotype supplies an adenovirus serotype lacking this gene (gene product) from another serotype with the missing function. For example, human adenovirus serotype 5 E1a and E1b functions are known to transcomplement E1-deleted chimpanzee adenovirus Pan 9. Similarly, the inventors have found that human Ad5 E1 transcomplements E1-deleted chimpanzee adenovirus serotypes Pan5, Pan6, Pan7, and simian adenovirus serotypes SV1, SV25 and SV39. Other examples of transcomplementing serotypes include human Ad5 and human Ad2, Ad3, Ad4, Ad5, Ad7, and Ad12.

[0019] The term "functionally deleted" or "functional deletion" means that a sufficient amount of the gene region

is removed or otherwise damaged, e.g., by mutation or modification, so that the gene region is no longer capable of producing functional products of gene expression. If desired, the entire gene region may be removed. Other suitable sites for gene disruption or deletion are discussed elsewhere in the application.

[0020] The term "functional" refers to a product (e.g., a protein or peptide) which performs its native function, although not necessarily at the same level as the native product. The term "functional" may also refer to a gene which encodes and from which a desired product can be expressed.

#### I. Chimeric Adenoviral Vectors

[0021] The compositions of this invention include chimeric adenoviral vectors that deliver a heterologous molecule to cells. For delivery of such a heterologous molecule, the vector can be a plasmid or, preferably, a chimeric adenovirus. The chimeric adenoviruses of the invention include adenovirus DNA from at least two source serotypes, a "donating serotype" and a "parental adenovirus" as described in more detail herein, and a minigene.

[0022] Because the adenoviral genome contains open reading frames on both strands, in many instances reference is made herein to 5' and 3' ends of the various regions to avoid confusion between specific open reading frames and gene regions. Thus, when reference is made herein to the "left" and "right" end of the adenoviral genome, this reference is to the ends of the approximately 36 kb adenoviral genome when depicted in schematic form as is conventional in the art [see, e.g., Horwitz, "Adenoviridae and Their Replication", in VIROLOGY, 2d ed., pp. 1679-1721 (1990)]. Thus, as used herein, the "left terminal end" of the adenoviral genome refers to portion of the adenoviral genome which, when the genome is depicted schematically in linear form, is located at the extreme left end of the schematic. Typically, the left end refers to be portion of the genome beginning at map unit 0 and extending to the right to include at least the 5' inverted terminal repeats (ITRs), and excludes the internal regions of the genome encoding the structural genes. As used herein, the "right terminal end" of the adenoviral genome refers to portion of the adenoviral genome which, when the genome is depicted schematically in linear form, is located at the extreme right end of the schematic. Typically, the right end of the adenoviral genome refers to be portion of the genome ending at map unit 36 and extending to the left to include at least the 3' ITRs, and excludes the internal regions of the genome encoding the structural genes.

#### [0023] A. Adenovirus Regulatory Sequences

##### [0024] 1. Serotype

[0025] The selection of the adenovirus serotype donating its left terminal end and right terminal end can be readily made by one of skill in the art from among serotypes which can readily be cultured in the desired cell line. Among other factors which may be considered in selecting the serotype of the donating serotype is compatibility with the adenovirus serotype which will be supplying the internal regions at the location at which their sequences are hybridized.

[0026] Suitable adenoviruses for donating their left and right termini are available from the American Type Culture

Collection, Manassas, Va., US (ATCC), a variety of academic and commercial sources, or the desired regions of the donating adenoviruses may be synthesized using known techniques with reference to sequences published in the literature or available from databases (e.g., GenBank, etc.). Examples of suitable donating adenoviruses include, without limitation, human adenovirus serotypes 2, 3, 4, 5, 7, and 12, and further including any of the presently identified human types [see, e.g., Horwitz, "Adenoviridae and Their Replication", in *VIROLOGY*, 2d ed., pp. 1679-1721 (1990)] which can be cultured in the desired cell. Similarly adenoviruses known to infect non-human primates (e.g., chimpanzees, rhesus, macaque, and other simian species) or other non-human mammals and which grow in the desired cell can be employed in the vector constructs of this invention. Such serotypes include, without limitation, chimpanzee adenoviruses Pan 5 [VR-591], Pan6 [VR-592], Pan7 [VR-593], and C68 (Pan9), described in U.S. Pat. No. 6,083,716; and simian adenoviruses including, without limitation SV1 [VR-195]; SV25 [SV-201]; SV35; SV15; SV-34; SV-36; SV-37, and baboon adenovirus [VR-275], among others. In the following examples, the human 293 cells and adenovirus type 5 (Ad5), Pan9, and Ad40 are used for convenience. However, one of skill in the art will understand that other cell lines and/or comparable regions derived from other adenoviral strains may be readily selected and used in the present invention in the place of (or in combination with) these serotypes.

#### [0027] 2. Sequences

[0028] The minimum sequences which must be supplied by the adenovirus donating its left terminal end and its right terminal end include the 5' cis-elements and the 3' cis-elements necessary for replication and packaging. Typically, the 5' cis-elements necessary for packaging and replication include the 5' inverted terminal repeat (ITR) sequences (which functions as origins of replication) and the native 5' packaging enhancer domains (that contain sequences necessary for packaging linear Ad genomes and enhancer elements for the E1 promoter). The right end of the adenoviral genome includes the 3' cis-elements (including the ITRs) necessary for packaging and encapsidation. Desirably, the adenovirus serotype donating its left and right termini and/or an adenovirus serotype which transcomplements the serotype of the donating adenovirus, further provides the functions of the necessary adenovirus early genes, including E1 (E1a and E1b), E2 (E2a and E2b), and E4 (including at least the ORF6 region). E3 is not essential and may be deleted as desired, e.g., for insertion of a transgene in this region or to provide space for a transgene inserted in another region (typically for packaging it is desirable for the total adenoviral genome to be under 36 kb).

[0029] In certain embodiments, the necessary adenovirus early genes are contained in the chimeric construct of the invention. In other embodiment, one or more of the necessary adenovirus early genes can be provided by the packaging host cell or in trans.

[0030] In general, the chimeric adenovirus of the invention contains regulatory sequences from the donating adenovirus serotype, or a transcomplementing serotype, to provide the chimeric adenovirus with compatible regulatory proteins. Optionally, one or more of the necessary adenoviral structural genes is provided by the adenovirus donating its left terminal and its right terminal end.

[0031] In certain embodiments, the chimeric adenovirus further contains one or more functional adenovirus genes, including, the Endoprotease open reading frame, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, protein VIII, pTP, 52/55 kDa protein, protein VII, Mu and/or protein VI from the adenovirus serotype donating its left and right termini. Where all of these genes are derived from the adenovirus serotype donating the 5' and 3' ITRs, a "pseudotyped" chimeric is formed. Optionally, one or more of the genes can be hybrids formed from the fusion of the donating adenovirus serotype and the parental adenovirus serotype providing the capsid proteins (e.g., without limitation, polymerase, terminal protein, IIIa protein). Suitably, these genes express functional proteins which permit packaging of the adenovirus genes into the capsid. Alternatively, one or more of these proteins (whether hybrid or non-hybrid) can be functionally deleted in the chimeric adenovirus. Where desired, any necessary proteins functionally deleted in the chimeric adenovirus can be expressed in trans in the packaging cell.

#### [0032] B. Parental Adenovirus Structural Proteins

##### [0033] 1. Serotypes

[0034] This invention is particularly well adapted for use in generating chimeric adenoviruses in which the capsid proteins are from a parental adenovirus which does not efficiently grow in a desirable host cell. The selection of the parental adenovirus serotype providing the internal regions can be readily made by one of skill in the art based on the information provided herein.

[0035] A variety of suitable adenoviruses can serve as a parental adenovirus supplying the regions encoding the structural (i.e., capsid proteins). Many of such adenoviruses can be obtained from the same sources as described above for the donating adenovirus serotypes. Examples of suitable parental adenovirus serotypes includes, without limitation, human adenovirus serotype 40, among others [see, e.g., Horwitz, "Adenoviridae and Their Replication", in *VIROLOGY*, 2d ed., pp. 1679-1721 (1990)], and adenoviruses known to infect non-human primates (e.g., chimpanzees, rhesus, macaque, and other simian species) or other non-human mammals, including, without limitation, chimpanzee adenovirus C1, described in U.S. Pat. No. 6,083,716; simian adenoviruses, and baboon adenoviruses, among others. In addition, the parental adenovirus supplying the internal regions may be from a non-naturally occurring adenovirus serotype, such as may be generated using a variety of techniques known to those of skill in the art.

[0036] In one embodiment illustrated herein, a chimeric virus that was constructed was that between the chimpanzee adenoviruses Pan-5 and C1 exhibited a higher titer in human 293 cells than the wild-type parental virus. However, the invention is not limited to the use of these chimpanzee adenoviruses, or to the combination of simian-simian, human-human, or simian-human chimeric adenoviruses. For example, it may be desirable to utilize bovine or canine adenoviruses, or other non-human mammalian adenoviruses which do not naturally infect and/or replicate in human cells.

[0037] In the following examples, the human adenovirus type 40 (Ad40) and the chimpanzee adenovirus C1, simian Pan 5 and Ad40, and Pan 5 and simian adenovirus SA18, are used. However, one of skill in the art will understand that

other adenoviral serotypes may be readily selected and used in the present invention in the place of (or in combination with) these serotypes.

**[0038]** 2. Sequences

**[0039]** The parental adenovirus provides to the chimeric construct of the invention its internal regions which includes structural proteins necessary for generating a capsid having the desired characteristics of the parental adenovirus. These desired characteristics include, but are not limited to, the ability to infect target cells and delivery a heterologous transgene, the ability to elude neutralizing antibodies directed to another adenovirus serotype (i.e., avoiding clearance due to cross-reactivity), and/or the ability to infect cells in the absence of an immune response to the chimeric adenovirus. The advantages of such characteristics may be most readily apparent in a regimen which involves repeat delivery of adenoviral vectors. The left and right termini of the parent adenovirus, including at least the 5' ITRs, the E1 region, the E4 region and the 3' ITRs are non-functional and, preferably, completely absent. Optionally, all adenovirus regulatory proteins from this parental adenovirus are non-functional and only the structural proteins (or selected structural proteins) are retained.

**[0040]** At a minimum, the parental adenovirus provides the adenoviral late region encoding the hexon protein. Suitably, the parental adenovirus further provides the late regions encoding the penton and the fiber. In certain embodiments, all of the functional adenoviral late regions, including L1 (encoding 52/55 Da, IIIa proteins), L2 (encoding penton, VII, V, Mu proteins), L3 (encoding VI, hexon, Endoprotease), L4 (encoding 100 kD, 33 kD, VIII proteins) and L5 (encoding fiber protein) are supplied by the parental adenovirus. Optionally, one or more of these late gene functions, with the exception of those encoding the hexon, penton and fiber proteins, can be functionally deleted. Any necessary structural proteins may be supplied in trans.

**[0041]** Thus, in certain embodiments, the chimeric adenovirus further contains one or more functional adenovirus genes, including, the Endoprotease open reading frame, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, protein VIII, pTP, 52/55 kDa protein, protein VII, Mu and/or protein VI from the parental adenovirus donating its internal regions. Optionally, one or more of the genes can be hybrids formed from the fusion of the donating adenovirus serotype and the parental adenovirus serotype providing the capsid proteins, as described above.

**[0042]** C. The "Minigene"

**[0043]** Typically, an adenoviral vector of the invention is designed to contain a minigene which may be inserted into the site of a partially deleted, fully deleted (absent), or disrupted adenoviral gene. For example, the minigene may be located in the site of such a functional E1 deletion or functional E3 deletion, or another suitable site.

**[0044]** The methods employed for the selection of the transgene, the cloning and construction of the "minigene" and its insertion into the viral vector are within the skill in the art given the teachings provided herein.

**[0045]** 1. The transgene

**[0046]** The transgene is a nucleic acid sequence, heterologous to the vector sequences flanking the transgene, which

encodes a polypeptide, protein, or other product, of interest. The nucleic acid coding sequence is operatively linked to regulatory components in a manner which permits transgene transcription, translation, and/or expression in a host cell.

**[0047]** The composition of the transgene sequence will depend upon the use to which the adenoviral vector will be put. For example, the adenoviral vector may be used as a helper virus in production of recombinant adeno-associated viruses or in production of recombinant adenoviruses deleted of essential adenoviral gene functions which are supplied by the adenoviral vector, or for a variety of production uses. Alternatively, the adenoviral vector may be used for diagnostic purposes.

**[0048]** One type of transgene sequence includes a reporter sequence, which upon expression produces a detectable signal. Such reporter sequences include, without limitation, DNA sequences encoding  $\beta$ -lactamase,  $\beta$ -galactosidase (LacZ), alkaline phosphatase, thymidine kinase, green fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT), luciferase, membrane bound proteins including, for example, CD2, CD4, CD8, the influenza hemagglutinin protein, and others well known in the art, to which high affinity antibodies directed thereto exist or can be produced by conventional means, and fusion proteins comprising a membrane bound protein appropriately fused to an antigen tag domain from, among others, hemagglutinin or Myc. These coding sequences, when associated with regulatory elements which drive their expression, provide signals detectable by conventional means, including enzymatic, radiographic, colorimetric, fluorescence or other spectrographic assays, fluorescent activating cell sorting assays and immunological assays, including enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and immunohistochemistry. For example, where the marker sequence is the LacZ gene, the presence of the vector carrying the signal is detected by assays for beta-galactosidase activity. Where the transgene is GFP or luciferase, the vector carrying the signal may be measured visually by color or light production in a luminometer.

**[0049]** However, desirably, the transgene is a non-marker sequence encoding a product which is useful in biology and medicine, such as proteins, peptides, RNA, enzymes, or catalytic RNAs. Desirable RNA molecules include tRNA, dsRNA, ribosomal RNA, si RNAs, small hairpin RNAs, trans-splicing RNAs, catalytic RNAs, and antisense RNAs. One example of a useful RNA sequence is a sequence which extinguishes expression of a targeted nucleic acid sequence in the treated animal.

**[0050]** The transgene may be used for treatment, e.g., of genetic deficiencies, as a cancer therapeutic or vaccine, for induction of an immune response, and/or for prophylactic vaccine purposes. As used herein, induction of an immune response refers to the ability of a molecule (e.g., a gene product) to induce a T cell and/or a humoral immune response to the molecule. The invention further includes using multiple transgenes, e.g., to correct or ameliorate a condition caused by a multi-subunit protein. In certain situations, a different transgene may be used to encode each subunit of a protein, or to encode different peptides or proteins. This is desirable when the size of the DNA encoding the protein subunit is large, e.g., for an immunoglobulin, the platelet-derived growth factor, or a dystrophin

protein. In order for the cell to produce the multi-subunit protein, a cell is infected with the recombinant virus containing each of the different subunits. Alternatively, different subunits of a protein may be encoded by the same transgene. In this case, a single transgene includes the DNA encoding each of the subunits, with the DNA for each subunit separated by an internal ribozyme entry site (IRES). This is desirable when the size of the DNA encoding each of the subunits is small, e.g., the total size of the DNA encoding the subunits and the IRES is less than five kilobases. As an alternative to an IRES, the DNA may be separated by sequences encoding a 2A peptide, which self-cleaves in a post-translational event. See, e.g., M.L. Donnelly, et al, *J. Gen. Virol.*, 78(Pt 1):13-21 (January 1997); Furler, S., et al, *Gene Ther.*, 8(11):864-873 (June 2001); Klump H., et al., *Gene Ther.*, 8(10):81 1-817 (May 2001). This 2A peptide is significantly smaller than an IRES, making it well suited for use when space is a limiting factor. However, the selected transgene may encode any biologically active product or other product, e.g., a product desirable for study.

[0051] Suitable transgenes may be readily selected by one of skill in the art. The selection of the transgene is not considered to be a limitation of this invention.

#### [0052] 2. Vector and Transgene Regulatory Elements

[0053] In addition to the major elements identified above for the minigene, the adenoviral vector also includes conventional control elements which are operably linked to the transgene in a manner that permits its transcription, translation and/or expression in a cell transfected with the plasmid vector or infected with the virus produced by the invention. As used herein, "operably linked" sequences include both expression control sequences that are contiguous with the gene of interest and expression control sequences that act in trans or at a distance to control the gene of interest.

[0054] Expression control sequences include appropriate transcription initiation, termination, promoter and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation (polyA) signals; sequences that stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (i.e., Kozak consensus sequence); sequences that enhance protein stability; and when desired, sequences that enhance secretion of the encoded product. A great number of expression control sequences, including promoters which are native, constitutive, inducible and/or tissue-specific, are known in the art and may be utilized.

[0055] Examples of constitutive promoters include, without limitation, the retroviral Rous sarcoma virus (RSV) LTR promoter (optionally with the RSV enhancer), the cytomegalovirus (CMV) promoter (optionally with the CMV enhancer) [see, e.g., Boshart et al, *Cell*, 41:521-530 (1985)], the SV40 promoter, the dihydrofolate reductase promoter, the  $\beta$ -actin promoter, the phosphoglycerol kinase (PGK) promoter, and the EF1a promoter [Invitrogen].

[0056] Inducible promoters allow regulation of gene expression and can be regulated by exogenously supplied compounds, environmental factors such as temperature, or the presence of a specific physiological state, e.g., acute phase, a particular differentiation state of the cell, or in replicating cells only. Inducible promoters and inducible systems are available from a variety of commercial sources,

including, without limitation, Invitrogen, Clontech and Ariad. Many other systems have been described and can be readily selected by one of skill in the art. For example, inducible promoters include the zinc-inducible sheep metallothionein (MT) promoter and the dexamethasone (Dex)-inducible mouse mammary tumor virus (MMTV) promoter. Other inducible systems include the T7 polymerase promoter system [WO 98/10088]; the ecdysone insect promoter [No et al, *Proc. Natl. Acad. Sci. USA*, 93:3346-3351 (1996)], the tetracycline-repressible system [Gossen et al, *Proc. Natl. Acad. Sci. USA*, 89:5547-5551 (1992)], the tetracycline-inducible system [Gossen et al, *Science*, 268:1766-1769 (1995), see also Harvey et al, *Curr. Opin. Chem. Biol.*, 2:512-518 (1998)]. Other systems include the FK506 dimer, VP16 or p65 using castradiol, diphenol murislerone, the RU486-inducible system [Wang et al, *Nat. Biotech.*, 15:239-243 (1997) and Wang et al, *Gene Ther.*, 4:432-441 (1997)] and the rapamycin-inducible system [Magari et al, *J. Clin. Invest.*, 100:2865-2872 (1997)]. The effectiveness of some inducible promoters increases over time. In such cases one can enhance the effectiveness of such systems by inserting multiple repressors in tandem, e.g., TetR linked to a TetR by an IRES. Alternatively, one can wait at least 3 days before screening for the desired function. One can enhance expression of desired proteins by known means to enhance the effectiveness of this system. For example, using the Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element (WPRE).

[0057] In another embodiment, the native promoter for the transgene will be used. The native promoter may be preferred when it is desired that expression of the transgene should mimic the native expression. The native promoter may be used when expression of the transgene must be regulated temporally or developmentally, or in a tissue-specific manner, or in response to specific transcriptional stimuli. In a further embodiment, other native expression control elements, such as enhancer elements, polyadenylation sites or Kozak consensus sequences may also be used to mimic the native expression.

[0058] Another embodiment of the transgene includes a transgene operably linked to a tissue-specific promoter. For instance, if expression in skeletal muscle is desired, a promoter active in muscle should be used. These include the promoters from genes encoding skeletal  $\beta$ -actin, myosin light chain 2A, dystrophin, muscle creatine kinase, as well as synthetic muscle promoters with activities higher than naturally occurring promoters (see Li et al., *Nat. Biotech.*, 17:241-245 (1999)). Examples of promoters that are tissue-specific are known for liver (albumin, Miyatake et al., *J. Virol.*, 71:5124-32 (1997); hepatitis B virus core promoter, Sandig et al., *Gene Ther.*, 3:1002-9 (1996); alpha-fetoprotein (AFP), Arbuthnot et al., *Hum. Gene Ther.*, 7:1503-14 (1996)), bone osteocalcin (Stein et al., *Mol. Biol. Rep.*, 24:185-96 (1997)); bone sialoprotein (Chen et al., *J. Bone Miner. Res.*, 11:654-64 (1996)), lymphocytes (CD2, Hansal et al., *J. Immunol.*, 161:1063-8 (1998); immunoglobulin heavy chain; T cell receptor chain), neuronal such as neuron-specific enolase (NSE) promoter (Andersen et al., *Cell. Mol. Neurobiol.*, 13:503-15 (1993)), neurofilament light-chain gene (Piccioli et al., *Proc. Natl. Acad. Sci. USA*, 88:5611-5 (1991)), and the neuron-specific vgf gene (Piccioli et al., *Neuron*, 15:373-84 (1995)), among others.

[0059] Optionally, vectors carrying transgenes encoding therapeutically useful or immunogenic products may also include selectable markers or reporter genes may include sequences encoding geneticin, hygromycin or purimycin resistance, among others. Such selectable reporters or marker genes (preferably located outside the viral genome to be packaged into a viral particle) can be used to signal the presence of the plasmids in bacterial cells, such as ampicillin resistance. Other components of the vector may include an origin of replication. Selection of these and other promoters and vector elements are conventional and many such sequences are available [see, e.g., Sambrook et al, and references cited therein].

[0060] These vectors are generated using the techniques and sequences provided herein, in conjunction with techniques known to those of skill in the art. Such techniques include conventional cloning techniques of cDNA such as those described in texts [Sambrook et al, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, N.Y.], use of overlapping oligonucleotide sequences of the adenovirus genomes, polymerase chain reaction, and any suitable method which provides the desired nucleotide sequence.

## II. Production of the Recombinant Viral Particle

[0061] In one embodiment, the invention provides a method of generating recombinant chimeric adenoviral particles in which the capsid of the chimeric adenovirus is of a serotype incapable of efficient growth in the selected host cell. A vector suitable for production of recombinant chimeric adenoviral particles can be generated homologous recombination between a first vector containing the left end of the chimeric adenoviral genome and a second vector containing the right end of the chimeric adenoviral genome. However, any suitable methodology known to those of skill in the art can be readily utilized to generate a vector suitable to generate a production vector, preferably which contains the entire chimeric adenoviral genome, including a mini-gene. This production vector is then introduced into a host cell in which the adenoviral capsid protein is assembled and the chimeric adenoviral particle assembled as described.

[0062] The chimeric adenoviruses of the invention include those in which one or more adenoviral genes are absent, or otherwise rendered non-functional. If any of the missing gene functions are essential to the replication and infectivity of the adenoviral particle, these functions are supplied by a complementation (or transcomplementing) cell line or a helper vector expressing these functions during production of the chimeric adenoviral particle.

[0063] Examples of chimeric adenoviruses containing such missing adenoviral gene functions include those which are partially or completely deleted in the E1a and/or E1b gene. In such a case, the E1 gene functions can be supplied by the packaging host cell, permitting the chimeric construct to be deleted of E1 gene functions and, if desired, for a transgene to be inserted in this region. Optionally, the E1 gene can be of a serotype which transcomplements the serotype providing the other adenovirus sequences in order to further reduce the possibility of recombination and improve safety. In other embodiments, it is desirable to retain an intact E1a and/or E1b region in the recombinant adenoviruses. Such an intact E1 region may be located in its

native location in the adenoviral genome or placed in the site of a deletion in the native adenoviral genome (e.g., in the E3 region).

[0064] In another example, all or a portion of the adenovirus delayed early gene E3 may be eliminated from the chimeric adenovirus. The function of adenovirus E3 is believed to be irrelevant to the function and production of the recombinant virus particle. Chimeric adenovirus vectors may also be constructed having a deletion of at least the ORF6 region of the E4 gene, and more desirably because of the redundancy in the function of this region, the entire E4 region. Still another vector of this invention contains a deletion in the delayed early gene E2a. Similarly, deletions in the intermediate genes IX and IVa<sub>2</sub> may be useful for some purposes. Optionally, deletions may also be made in selected portions of the late genes L1 through L5, as described above.

[0065] Other deletions may be made in the other structural or non-structural adenovirus genes. The above-discussed deletions may be used individually, i.e., an adenovirus sequence for use in the present invention may contain deletions in only a single region. Alternatively, deletions of entire genes or portions thereof effective to destroy their biological activity may be used in any combination. For example, in one exemplary vector, the adenovirus sequence may have deletions of the E1 genes and the E4 gene, or of the E1, E2a and E3 genes, or of the E1 and E3 genes, or of E1, E2a and E4 genes, with or without deletion of E3, and so on. As discussed above, such deletions may be used in combination with other mutations, such as temperature-sensitive mutations, to achieve a desired result.

[0066] Examples of suitable transcomplementing serotypes are provided above. The use of transcomplementing serotypes can be particularly advantageous where there is diversity between the Ad sequences in the vector of the invention and the human AdE1 sequences found in currently available packaging cells. In such cases, the use of the current human E1-containing cells prevents the generation of replication-competent adenoviruses during the replication and production process. However, in certain circumstances, it will be desirable to utilize a cell line which expresses the E1 gene products can be utilized for production of an E1-deleted simian adenovirus. Such cell lines have been described. See, e.g., U.S. Pat. No. 6,083,716.

### [0067] A. Packaging Host Cells

[0068] Suitably, the packaging host cell is selected from among cells in which the adenovirus serotype donating the left and right terminal ends of the chimeric genome are capable of efficient growth. The host cells are preferably of mammalian origin, and most preferably are of non-human primate or human origin.

[0069] Particularly desirable host cells are selected from among any mammalian species, including, without limitation, cells such as A549 [ATCC Accession No. CCL 185], 911 cells, WEHI, 3T3, 10T1/2, HEK 293 cells or PERC6 (both of which express functional adenoviral E1) [Fallaux, F J et al, (1998), *Hum Gene Ther*, 9:1909-1917], Saos, C2C12, L cells, HT1080, HepG2, HeLa [ATCC Accession No. CCL 2], KB [CCL 17], Detroit [e.g., Detroit 510, CCL 72] and WI-38 [CCL 75] cells, and primary fibroblast, hepatocyte and myoblast cells derived from mammals

including human, monkey, mouse, rat, rabbit, and hamster. These cell lines are all available from the American Type Culture Collection, 10801 University Boulevard, Manassas, Va. 20110-2209. Other suitable cell lines may be obtained from other sources. The selection of the mammalian species providing the cells is not a limitation of this invention; nor is the type of mammalian cell, i.e., fibroblast, hepatocyte, tumor cell, etc.

[0070] As described above, a chimeric adenovirus of the invention can lack one or more functional adenoviral regulatory and/or structural genes which are supplied either by the host cell or in trans to effect packaging of the chimeric adenovirus into the viral capsid to generate the viral particle. Thus, the ability of a selected host cell to supply transcomplementing adenoviral sequences may be taken into consideration in selecting a desired host cell.

[0071] In one example, the cells are from a stable cell line which expresses adenovirus E1a and E1b functions from a cell line which transcomplements the adenovirus serotype which donates the left and right termini to the chimera of the invention, permitting the chimera to be E1-deleted. Alternatively, where the cell line does not transcomplement the adenovirus donating the termini, E1 functions may be provided by the chimera, or in trans.

[0072] If desired, one may utilize the sequences provided herein to generate a packaging cell or cell line that expresses, at a minimum, the adenovirus E1 gene from the adenovirus serotype donating the 5' ITR under the transcriptional control of a promoter for expression, or a transcomplementing serotype, in a selected parent cell line. Inducible or constitutive promoters may be employed for this purpose. Examples of such promoters are described in detail elsewhere in this specification. A parent cell is selected for the generation of a novel cell line expressing any desired adenovirus or adenovirus gene, including, e.g., a human Ad5, AdPan5, Pan6, Pan7, SV1, SV25 or SV39 gene. Without limitation, such a parent cell line may be HeLa [ATCC Accession No. CCL 2], A549 [ATCC Accession No. CCL 185], HEK 293, KB [CCL 17], Detroit [e.g., Detroit 510, CCL 72] and WI-38 [CCL 75] cells, among others. Many of these cell lines are all available from the ATCC. Other suitable parent cell lines may be obtained from other sources.

[0073] Such E1-expressing cell lines are useful in the generation of chimeric adenovirus E1 deleted vectors. Additionally, or alternatively, the invention provides cell lines that express one or more simian adenoviral gene products, e.g., E1a, E1b, E2a, and/or E4 ORF6, can be constructed using essentially the same procedures for use in the generation of chimeric viral vectors. Such cell lines can be utilized to transcomplement adenovirus vectors deleted in the essential genes that encode those products, or to provide helper functions necessary for packaging of a helper-dependent virus (e.g., adeno-associated virus). The preparation of a host cell according to this invention involves techniques such as assembly of selected DNA sequences. This assembly may be accomplished utilizing conventional techniques. Such techniques include cDNA and genomic cloning, which are well known and are described in Sambrook et al., cited above, use of overlapping oligonucleotide sequences of the adenovirus genomes, combined with polymerase chain reaction, synthetic methods, and any other suitable methods which provide the desired nucleotide sequence.

[0074] In still another alternative, the essential adenoviral gene products are provided in trans by the adenoviral vector and/or helper virus. In such an instance, a suitable host cell can be selected from any biological organism, including prokaryotic (e.g., bacterial) cells, and eukaryotic cells, including, insect cells, yeast cells and mammalian cells. Particularly desirable host cells are selected from among any mammalian species, including, without limitation, cells such as A549, WEHI, 3T3, 10T1/2, HEK 293 cells or PERC6 (both of which express functional adenoviral E1) [Fallaux, F J et al, (1998), Hum Gene Ther, 9:1909-1917], Saos, C2C12, L cells, HT1080, HepG2 and primary fibroblast, hepatocyte and myoblast cells derived from mammals including human, monkey, mouse, rat, rabbit, and hamster. The selection of the mammalian species providing the cells is not a limitation of this invention; nor is the type of mammalian cell, i.e., fibroblast, hepatocyte, tumor cell, etc.

#### [0075] B. Helper Vectors

[0076] Thus, depending upon the adenovirus gene content of the adenoviral vectors and any adenoviral gene functions expressed from the host cell, a helper vector may be necessary to provide sufficient adenovirus gene sequences necessary to produce an infective recombinant viral particle containing the minigene. See, for example, the techniques described for preparation of a "minimal" human Ad vector in International Patent Application WO96/13597, published May 9, 1996, and incorporated herein by reference. Suitably, these helper vectors may be non-replicating genetic elements, a plasmid, or a virus.

[0077] Useful helper vectors contain selected adenovirus gene sequences not present in the adenovirus vector construct and/or not expressed by the packaging cell line in which the vector is transfected. In one embodiment, the helper virus is replication-defective and contains a variety of adenovirus genes in addition to the sequences described above. Such a helper vector is desirably used in combination with an E1-expressing cell line.

[0078] Helper vectors may be formed into poly-cation conjugates as described in Wu et al, *J. Biol. Chem.*, 264:16985-16987 (1989); K. J. Fisher and J. M. Wilson, *Biochem. J.*, 299:49 (Apr. 1, 1994). A helper vector may optionally contain a second reporter minigene. A number of such reporter genes are known to the art. The presence of a reporter gene on the helper virus which is different from the transgene on the adenovirus vector allows both the Ad vector and the helper vector to be independently monitored. This second reporter is used to enable separation between the resulting recombinant virus and the helper virus upon purification.

#### [0079] C. Assembly of Viral Particle and Transfection of a Cell Line

[0080] Generally, when delivering the vector comprising the minigene by transfection, the vector is delivered in an amount from about 5  $\mu\text{g}$  to about 100  $\mu\text{g}$  DNA, and preferably about 10 to about 50  $\mu\text{g}$  DNA to about  $1 \times 10^4$  cells to about  $1 \times 10^3$  cells, and preferably about  $10^5$  cells. However, the relative amounts of vector DNA to host cells may be adjusted, taking into consideration such factors as the selected vector, the delivery method and the host cells selected.

[0081] Introduction into the host cell of the vector may be achieved by any means known in the art or as disclosed



above, including transfection, and infection. One or more of the adenoviral genes may be stably integrated into the genome of the host cell, stably expressed as episomes, or expressed transiently. The gene products may all be expressed transiently, on an episome or stably integrated, or some of the gene products may be expressed stably while others are expressed transiently.

[0082] Furthermore, the promoters for each of the adenoviral genes may be selected independently from a constitutive promoter, an inducible promoter or a native adenoviral promoter. The promoters may be regulated by a specific physiological state of the organism or cell (i.e., by the differentiation state or in replicating or quiescent cells) or by exogenously added factors, for example.

[0083] Introduction of the molecules (as plasmids or viruses) into the host cell may also be accomplished using techniques known to the skilled artisan and as discussed throughout the specification. In preferred embodiment, standard transfection techniques are used, e.g., CaPO<sub>4</sub> transfection or electroporation.

[0084] Assembly of the selected DNA sequences of the adenovirus (as well as the transgene and other vector elements) into various intermediate plasmids, and the use of the plasmids and vectors to produce a recombinant viral particle are all achieved using conventional techniques. Such techniques include conventional cloning techniques of cDNA such as those described in texts [Sambrook et al, cited above], use of overlapping oligonucleotide sequences of the adenovirus genomes, polymerase chain reaction, and any suitable method which provides the desired nucleotide sequence. Standard transfection and co-transfection techniques are employed, e.g., CaPO<sub>4</sub> precipitation techniques. Other conventional methods employed include homologous recombination of the viral genomes, plaquing of viruses in agar overlay, methods of measuring signal generation, and the like.

[0085] For example, following the construction and assembly of the desired minigene-containing viral vector, the vector is transfected in vitro in the presence of an optional helper vector into the packaging cell line. The functions expressed from the plasmid, packaging cell line and helper virus, if any, permits the adenovirus-transgene sequences in the vector to be replicated and packaged into virion capsids, resulting in the chimeric viral particles. The current method for producing such virus particles is transfection-based. However, the invention is not limited to such methods. The resulting chimeric adenoviruses are useful in transferring a selected transgene to a selected cell.

### III. Use of the Chimeric Adenovirus Vectors

[0086] The chimeric adenovirus vectors of the invention are useful for gene transfer to a human or veterinary subject (including, non-human primates, non-simian primates, and other mammals) in vitro, ex vivo, and in vivo.

[0087] The recombinant adenovirus vectors described herein can be used as expression vectors for the production of the products encoded by the heterologous genes in vitro. For example, the recombinant adenoviruses containing a gene inserted into the location of an E1 deletion may be transfected into an E1-expressing cell line as described above. Alternatively, replication-competent adenoviruses may be used in another selected cell line. The transfected

cells are then cultured in the conventional manner, allowing the recombinant adenovirus to express the gene product from the promoter. The gene product may then be recovered from the culture medium by known conventional methods of protein isolation and recovery from culture.

[0088] A chimeric adenoviral vector of the invention provides an efficient gene transfer vehicle that can deliver a selected transgene to a selected host cell in vivo or ex vivo even where the organism has neutralizing antibodies to one or more AAV serotypes. In one embodiment, the rAAV and the cells are mixed ex vivo; the infected cells are cultured using conventional methodologies; and the transduced cells are re-infused into the patient. These compositions are particularly well suited to gene delivery for therapeutic purposes and for immunization, including inducing protective immunity.

[0089] More commonly, the chimeric adenoviral vectors of the invention will be utilized for delivery of therapeutic or immunogenic molecules, as described below. It will be readily understood for both applications that the recombinant adenoviral vectors of the invention are particularly well suited for use in regimens involving repeat delivery of recombinant adenoviral vectors. Such regimens typically involve delivery of a series of viral vectors in which the viral capsids are alternated. The viral capsids may be changed for each subsequent administration, or after a pre-selected number of administrations of a particular serotype capsid (e.g., one, two, three, four or more). Thus, a regimen may involve delivery of a rAd with a first capsid, delivery with a rAd with a second capsid, and delivery with a third capsid. A variety of other regimens which use the Ad capsids of the invention alone, in combination with one another, or in combination with other Ad serotypes will be apparent to those of skill in the art. Optionally, such a regimen may involve administration of rAd with capsids of non-human primate adenoviruses, human adenoviruses, or artificial (e.g., chimeric) serotypes such as are described herein. Each phase of the regimen may involve administration of a series of injections (or other delivery routes) with a single Ad serotype capsid followed by a series with another Ad serotype capsid. Alternatively, the recombinant Ad vectors of the invention may be utilized in regimens involving other non-adenoviral-mediated delivery systems, including other viral systems, non-viral delivery systems, protein, peptides, and other biologically active molecules.

[0090] The following sections will focus on exemplary molecules which may be delivered via the adenoviral vectors of the invention.

#### [0091] A. Ad-Mediated Delivery of Therapeutic Molecules

[0092] In one embodiment, the Ad vectors described herein are administered to humans according to published methods for gene therapy. A viral vector of the invention bearing the selected transgene may be administered to a patient, preferably suspended in a biologically compatible solution or pharmaceutically acceptable delivery vehicle. A suitable vehicle includes sterile saline. Other aqueous and non-aqueous isotonic sterile injection solutions and aqueous and non-aqueous sterile suspensions known to be pharmaceutically acceptable carriers and well known to those of skill in the art may be employed for this purpose.

[0093] The adenoviral vectors are administered in sufficient amounts to transduce the target cells and to provide

sufficient levels of gene transfer and expression to provide a therapeutic benefit without undue adverse or with medically acceptable physiological effects, which can be determined by those skilled in the medical arts. Conventional and pharmaceutically acceptable routes of administration include, but are not limited to, direct delivery to the retina and other intraocular delivery methods, direct delivery to the liver, inhalation, intranasal, intravenous, intramuscular, intratracheal, subcutaneous, intradermal, rectal, oral and other parenteral routes of administration. Routes of administration may be combined, if desired, or adjusted depending upon the transgene or the condition. The route of administration primarily will depend on the nature of the condition being treated.

[0094] Dosages of the viral vector will depend primarily on factors such as the condition being treated, the age, weight and health of the patient, and may thus vary among patients. For example, a therapeutically effective adult human or veterinary dosage of the viral vector is generally in the range of from about 100  $\mu$ L to about 100 mL of a carrier containing concentrations of from about  $1 \times 10^6$  to about  $1 \times 10^{15}$  particles, about  $1 \times 10^{11}$  to  $1 \times 10^{13}$  particles, or about  $1 \times 10^9$  to  $1 \times 10^{12}$  particles virus. Dosages will range depending upon the size of the animal and the route of administration. For example, a suitable human or veterinary dosage (for about an 80 kg animal) for intramuscular injection is in the range of about  $1 \times 10^9$  to about  $5 \times 10^{12}$  particles per mL, for a single site. Optionally, multiple sites of administration may be delivered. In another example, a suitable human or veterinary dosage may be in the range of about  $1 \times 10^{11}$  to about  $1 \times 10^{15}$  particles for an oral formulation. One of skill in the art may adjust these doses, depending the route of administration, and the therapeutic or vaccinal application for which the recombinant vector is employed. The levels of expression of the transgene, or for an immunogen, the level of circulating antibody, can be monitored to determine the frequency of dosage administration. Yet other methods for determining the timing of frequency of administration will be readily apparent to one of skill in the art.

[0095] An optional method step involves the co-administration to the patient, either concurrently with, or before or after administration of the viral vector, of a suitable amount of a short acting immune modulator. The selected immune modulator is defined herein as an agent capable of inhibiting the formation of neutralizing antibodies directed against the recombinant vector of this invention or capable of inhibiting cytolytic T lymphocyte (CTL) elimination of the vector. The immune modulator may interfere with the interactions between the T helper subsets ( $T_{H1}$  or  $T_{H2}$ ) and B cells to inhibit neutralizing antibody formation. Alternatively, the immune modulator may inhibit the interaction between  $T_{H1}$  cells and CTLs to reduce the occurrence of CTL elimination of the vector. A variety of useful immune modulators and dosages for use of same are disclosed, for example, in Yang et al., *J. Virol.*, 70(9) (September 1996); International Patent Application No. WO96/12406, published May 2, 1996; and International Patent Application No. PCT/US96/03035, all incorporated herein by reference. Typically, such immune modulators would be selected when the transgene is a therapeutic which requires repeat delivery.

#### [0096] 1. Therapeutic Transgenes

[0097] Useful therapeutic products encoded by the transgene include hormones and growth and differentiation factors including, without limitation, insulin, glucagon, growth hormone (GH), parathyroid hormone (PTH), growth hormone releasing factor (GRF), follicle stimulating hormone (FSH), luteinizing hormone (LH), human chorionic gonadotropin (hCG), vascular endothelial growth factor (VEGF), angiopoietins, angiostatin, granulocyte colony stimulating factor (GCSF), erythropoietin (EPO), connective tissue growth factor (CTGF), basic fibroblast growth factor (bFGF), acidic fibroblast growth factor (aFGF), epidermal growth factor (EGF), transforming growth factor  $\alpha$  (TGF  $\alpha$ ), platelet-derived growth factor (PDGF), insulin growth factors I and II (IGF-I and IGF-II), any one of the transforming growth factor superfamily, including TGF, activins, inhibins, or any of the bone morphogenic proteins (BMP) BMPs 1-15, any one of the heregluin/neuregulin/ARIA/neu differentiation factor (NDF) family of growth factors, nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), neurotrophins NT-3 and NT-4/5, ciliary neurotrophic factor (CNTF), glial cell line derived neurotrophic factor (GDNF), neurturin, agrin, any one of the family of semaphorins/collapsins, netrin-1 and netrin-2, hepatocyte growth factor (HGF), ephrins, noggin, sonic hedgehog and tyrosine hydroxylase.

[0098] Other useful transgene products include proteins that regulate the immune system including, without limitation, cytokines and lymphokines such as thrombopoietin (TPO), interleukins (IL) IL-1 through IL-25 (including, e.g., IL-2, IL-4, IL -12 and IL-18), monocyte chemoattractant protein, leukemia inhibitory factor, granulocyte-macrophage colony stimulating factor, Fas ligand, tumor necrosis factors and, interferons, and, stem cell factor, flk-2/flt3 ligand. Gene products produced by the immune system are also useful in the invention. These include, without limitation, immunoglobulins IgG, IgM, IgA, IgD and IgE, chimeric immunoglobulins, humanized antibodies, single chain antibodies, T cell receptors, chimeric T cell receptors, single chain T cell receptors, class I and class II MHC molecules, as well as engineered immunoglobulins and MHC molecules. Useful gene products also include complement regulatory proteins such as complement regulatory proteins, membrane cofactor protein (MCP), decay accelerating factor (DAF), CRI, CF2 and CD59.

[0099] Still other useful gene products include any one of the receptors for the hormones, growth factors, cytokines, lymphokines, regulatory proteins and immune system proteins. The invention encompasses receptors for cholesterol regulation, including the low density lipoprotein (LDL) receptor, high density lipoprotein (HDL) receptor, the very low density lipoprotein (VLDL) receptor, proteins useful in the regulation of lipids, including, e.g., apolipoprotein (apo) A and its isoforms (e.g., ApoA1), apoE and its isoforms including E2, E3 and E4), SRB1, ABC1, and the scavenger receptor. The invention also encompasses gene products such as members of the steroid hormone receptor superfamily including glucocorticoid receptors and estrogen receptors, Vitamin D receptors and other nuclear receptors. In addition, useful gene products include transcription factors such as jun, fos, max, mad, serum response factor (SRF), AP-1, AP2, myb, MyoD and myogenin, ETS-box containing proteins, TFE3, E2F, ATF1, ATF2, ATF3, ATF4, ZF5, NFAT,

CREB, HNF-4, C/EBP, SP1, CCAAT-box binding proteins, interferon regulation factor (IRF-1), Wilms tumor protein, ETS-binding protein, STAT, GATA-box binding proteins, e.g., GATA-3, and the forkhead family of winged helix proteins.

[0100] Other useful gene products include, carbamoyl synthetase I, ornithine transcarbamylase, arginosuccinate synthetase, arginosuccinate lyase, arginase, fumarylacetylacetate hydrolase, phenylalanine hydroxylase, alpha-1 antitrypsin, glucose-6-phosphatase, porphobilinogen deaminase, cystathione beta-synthase, branched chain ketoacid decarboxylase, albumin, isovaleryl-coA dehydrogenase, propionyl CoA carboxylase, methyl malonyl CoA mutase, glutaryl CoA dehydrogenase, insulin, beta-glucosidase, pyruvate carboxylate, hepatic phosphorylase, phosphorylase kinase, glycine decarboxylase, H-protein, T-protein, a cystic fibrosis transmembrane regulator (CFTR) sequence, and a dystrophin cDNA sequence. Other useful gene products include those useful for treatment of hemophilia A (e.g., Factor VIII and its variants, including the light chain and heavy chain of the heterodimer, optionally operably linked by a junction), and the B-domain deleted Factor VIII, see U.S. Pat. Nos. 6,200,560 and 6,221,349], and useful for treatment of hemophilia B (e.g., Factor IX).

[0101] Still other useful gene products include non-naturally occurring polypeptides, such as chimeric or hybrid polypeptides having a non-naturally occurring amino acid sequence containing insertions, deletions or amino acid substitutions. For example, single-chain engineered immunoglobulins could be useful in certain immunocompromised patients. Other types of non-naturally occurring gene sequences include antisense molecules and catalytic nucleic acids, such as ribozymes, which could be used to reduce overexpression of a target.

[0102] Reduction and/or modulation of expression of a gene are particularly desirable for treatment of hyperproliferative conditions characterized by hyperproliferating cells, as are cancers and psoriasis. Target polypeptides include those polypeptides which are produced exclusively or at higher levels in hyperproliferative cells as compared to normal cells. Target antigens include polypeptides encoded by oncogenes such as myb, myc, fyn, and the translocation gene bcr/abl, ras, src, P53, neu, trk and EGRF. In addition to oncogene products as target antigens, target polypeptides for anti-cancer treatments and protective regimens include variable regions of antibodies made by B cell lymphomas and variable regions of T cell receptors of T cell lymphomas which, in some embodiments, are also used as target antigens for autoimmune disease. Other tumor-associated polypeptides can be used as target polypeptides such as polypeptides which are found at higher levels in tumor cells including the polypeptide recognized by monoclonal antibody 17-1A and folate binding polypeptides.

[0103] Other suitable therapeutic polypeptides and proteins include those which may be useful for treating individuals suffering from autoimmune diseases and disorders by conferring a broad based protective immune response against targets that are associated with autoimmunity including cell receptors and cells which produce self-directed antibodies. T-cell mediated autoimmune diseases include rheumatoid arthritis (RA), multiple sclerosis (MS), Sjogren's syndrome, sarcoidosis, insulin dependent diabetes

mellitus (IDDM), autoimmune thyroiditis, reactive arthritis, ankylosing spondylitis, scleroderma, polymyositis, dermatomyositis, psoriasis, vasculitis, Wegener's granulomatosis, Crohn's disease and ulcerative colitis. Each of these diseases is characterized by T cell receptors (TCRs) that bind to endogenous antigens and initiate the inflammatory cascade associated with autoimmune diseases.

[0104] The chimeric adenoviral vectors of the invention are particularly well suited for therapeutic regimens in which multiple adenoviral-mediated deliveries of transgenes is desired, e.g., in regimens involving redelivery of the same transgene or in combination regimens involving delivery of other transgenes. Such regimens may involve administration of a chimeric adenoviral vector, followed by re-administration with a vector from the same serotype adenovirus. Particularly desirable regimens involve administration of a chimeric adenoviral vector of the invention, in which the serotype of the viral vector delivered in the first administration differs from the serotype of the viral vector utilized in one or more of the subsequent administrations. For example, a therapeutic regimen involves administration of a chimeric vector and repeat administration with one or more adenoviral vectors of the same or different serotypes. In another example, a therapeutic regimen involves administration of an adenoviral vector followed by repeat administration with a chimeric vector of the invention which differs from the serotype of the first delivered adenoviral vector, and optionally further administration with another vector which is the same or, preferably, differs from the serotype of the vector in the prior administration steps. These regimens are not limited to delivery of adenoviral vectors constructed using the chimeric serotypes of the invention. Rather, these regimens can readily utilize chimeric or non-chimeric vectors of other adenoviral serotypes, which may be of artificial, human or non-human primate, or other mammalian sources, in combination with one or more of the chimeric vectors of the invention. Examples of such serotypes are discussed elsewhere in this document. Further, these therapeutic regimens may involve either simultaneous or sequential delivery of chimeric adenoviral vectors of the invention in combination with non-adenoviral vectors, non-viral vectors, and/or a variety of other therapeutically useful compounds or molecules. The present invention is not limited to these therapeutic regimens, a variety of which will be readily apparent to one of skill in the art.

[0105] B. Ad-Mediated Delivery of Immunogenic Transgenes

[0106] The adenoviruses of the invention may also be employed as immunogenic compositions. As used herein, an immunogenic composition is a composition to which a humoral (e.g., antibody) or cellular (e.g., a cytotoxic T cell) response is mounted to a transgene product delivered by the immunogenic composition following delivery to a mammal, and preferably a primate. The present invention provides an Ad that can contain in any of its adenovirus sequence deletions a gene encoding a desired immunogen. Chimeric adenoviruses based on simian or other non-human mammalian primate serotypes are likely to be better suited for use as a live recombinant virus vaccine in different animal species compared to an adenovirus of human origin, but is not limited to such a use. The recombinant adenoviruses can be used as prophylactic or therapeutic vaccines against any pathogen for which the antigen(s) crucial for induction of an

immune response and able to limit the spread of the pathogen has been identified and for which the CDNA is available.

[0107] Such vaccinal (or other immunogenic) compositions are formulated in a suitable delivery vehicle, as described above. Generally, doses for the immunogenic compositions are in the range defined above for therapeutic compositions. The levels of immunity of the selected gene can be monitored to determine the need, if any, for boosters. Following an assessment of antibody titers in the serum, optional booster immunizations may be desired.

[0108] Optionally, a vaccinal composition of the invention may be formulated to contain other components, including, e.g. adjuvants, stabilizers, pH adjusters, preservatives and the like. Such components are well known to those of skill in the vaccine art. Examples of suitable adjuvants include, without limitation, liposomes, alum, monophosphoryl lipid A, and any biologically active factor, such as cytokine, an interleukin, a chemokine, a ligands, and optimally combinations thereof. Certain of these biologically active factors can be expressed *in vivo*, e.g., via a plasmid or viral vector. For example, such an adjuvant can be administered with a priming DNA vaccine encoding an antigen to enhance the antigen-specific immune response compared with the immune response generated upon priming with a DNA vaccine encoding the antigen only.

[0109] The adenoviruses are administered in "an immunogenic amount", that is, an amount of adenovirus that is effective in a route of administration to transfect the desired cells and provide sufficient levels of expression of the selected gene to induce an immune response. Where protective immunity is provided, the recombinant adenoviruses are considered to be vaccine compositions useful in preventing infection and/or recurrent disease.

[0110] Alternatively, or in addition, the vectors of the invention may contain a transgene encoding a peptide, polypeptide or protein which induces an immune response to a selected immunogen. The recombinant adenoviruses of this invention are expected to be highly efficacious at inducing cytolytic T cells and antibodies to the inserted heterologous antigenic protein expressed by the vector.

[0111] For example, immunogens may be selected from a variety of viral families. Example of desirable viral families against which an immune response would be desirable include, the picornavirus family, which includes the genera rhinoviruses, which are responsible for about 50% of cases of the common cold; the genera enteroviruses, which include polioviruses, coxsackieviruses, echoviruses, and human enteroviruses such as hepatitis A virus; and the genera aphoviruses, which are responsible for foot and mouth diseases, primarily in non-human animals. Within the picornavirus family of viruses, target antigens include the VP1, VP2, VP3, VP4, and VPG. Another viral family includes the calcivirus family, which encompasses the Norwalk group of viruses, which are an important causative agent of epidemic gastroenteritis. Still another viral family desirable for use in targeting antigens for inducing immune responses in humans and non-human animals is the togavirus family, which includes the genera alphavirus, which include Sindbis viruses, Ross River virus, and Venezuelan, Eastern & Western Equine encephalitis, and rubivirus, including Rubella virus. The flaviviridae family includes

dengue, yellow fever, Japanese encephalitis, St. Louis encephalitis and tick borne encephalitis viruses. Other target antigens may be generated from the Hepatitis C or the coronavirus family, which includes a number of non-human viruses such as infectious bronchitis virus (poultry), porcine transmissible gastroenteric virus (pig), porcine hemagglutinin encephalomyelitis virus (pig), feline infectious peritonitis virus (cats), feline enteric coronavirus (cat), canine coronavirus (dog), and human respiratory coronaviruses, which may cause the common cold and/or non-A, B or C hepatitis. In addition, the human coronaviruses include the putative causative agent of sudden acute respiratory syndrome (SARS). Within the coronavirus family, target antigens include the E1 (also called M or matrix protein), E2 (also called S or Spike protein), E3 (also called HE or hemagglutinin-esterase) glycoprotein (not present in all coronaviruses), or N (nucleocapsid). Still other antigens may be targeted against the rhabdovirus family, which includes the genera vesiculovirus (e.g., Vesicular Stomatitis Virus), and the general lyssavirus (e.g., rabies). Within the rhabdovirus family, suitable antigens may be derived from the G protein or the N protein. The family filoviridae, which includes hemorrhagic fever viruses such as Marburg and Ebola virus, may be a suitable source of antigens. The paramyxovirus family includes parainfluenza Virus Type 1, parainfluenza Virus Type 3, bovine parainfluenza Virus Type 3, rubulavirus (mumps virus), parainfluenza Virus Type 2, parainfluenza virus Type 4, Newcastle disease virus (chickens), rinderpest, morbillivirus, which includes measles and canine distemper, and pneumovirus, which includes respiratory syncytial virus. The influenza virus is classified within the family orthomyxovirus and is a suitable source of antigen (e.g., the HA protein, the N1 protein). The bunyavirus family includes the genera bunyavirus (California encephalitis, La Crosse), phlebovirus (Rift Valley Fever), hantavirus (pneumonia is a hemorrhagic fever virus),airovirus (Nairobi sheep disease) and various unassigned bunyaviruses. The arenavirus family provides a source of antigens against LCM and Lassa fever virus. The reovirus family includes the genera reovirus, rotavirus (which causes acute gastroenteritis in children), orbiviruses, and cultivirus (Colorado Tick fever), Leishmania (humans), equine encephalosis, blue tongue.

[0112] The retrovirus family includes the sub-family oncovirinal which encompasses such human and veterinary diseases as feline leukemia virus, HTLV I and HTLV II, lentivirinal (which includes human immunodeficiency virus (HIV), simian immunodeficiency virus (SIV), feline immunodeficiency virus (FIV), equine infectious anemia virus, and spumavirinal). Among the lentiviruses, many suitable antigens have been described and can readily be selected. Examples of suitable HIV and SIV antigens include, without limitation the gag, pol, Vif, Vpx, VPR, Env, Tat, Nef, and Rev proteins, as well as various fragments thereof. For example, suitable fragments of the Env protein may include any of its subunits such as the gp120, gp160, gp41, or smaller fragments thereof, e.g., of at least about 8 amino acids in length. Similarly, fragments of the tat protein may be selected. [See, U.S. Pat. Nos. 5,891,994 and 6,193,981.] See, also, the HIV and SIV proteins described in D. H. Barouch et al, *J. Virol.*, 75(5):2462-2467 (March 2001), and R. R. Amara, et al, *Science*, 292:69-74 (6 April 2001). In another example, the HIV and/or SIV immunogenic proteins or peptides may be used to form fusion proteins or other immunogenic molecules. See, e.g., the HIV-1 Tat and/or Nef

fusion proteins and immunization regimens described in WO 01/54719, published Aug. 2, 2001, and WO 99/16884, published Apr. 8, 1999. The invention is not limited to the HIV and/or SIV immunogenic proteins or peptides described herein. In addition, a variety of modifications to these proteins have been described or could readily be made by one of skill in the art. See, e.g., the modified gag protein that is described in U.S. Pat. No. 5,972,596. Further, any desired HIV and/or SIV immunogens may be delivered alone or in combination. Such combinations may include expression from a single vector or from multiple vectors. Optionally, another combination may involve delivery of one or more expressed immunogens with delivery of one or more of the immunogens in protein form. Such combinations are discussed in more detail below.

[0113] The papovavirus family includes the sub-family polyomaviruses (BKU and JCU viruses) and the sub-family papillomavirus (associated with cancers or malignant progression of papilloma). The adenovirus family includes viruses (EX, AD7, ARD, O.B.) which cause respiratory disease and/or enteritis. The parvovirus includes family feline parvovirus (feline enteritis), feline pantecopeniavirus, canine parvovirus, and porcine parvovirus. The herpesvirus family includes the sub-family alphaherpesvirinae, which encompasses the genera simplexvirus (HSV1, HSV2), varicellovirus (pseudorabies, varicella zoster) and the sub-family betaherpesvirinae, which includes the genera cytomegalovirus (HCMV, muromegalovirus) and the sub-family gammaherpesvirinae, which includes the genera lymphocryptovirus, EBV (Burkitts lymphoma), infectious rhinotracheitis, Marek's disease virus, and rhadinovirus. The poxvirus family includes the sub-family chordopoxvirinae, which encompasses the genera orthopoxvirus (Variola (Smallpox) and Vaccinia (Cowpox)), parapoxvirus, avipoxvirus, capripoxvirus, leporipoxvirus, suipoxvirus, and the sub-family entomopoxvirinae. The hepadnavirus family includes the Hepatitis B virus. One unclassified virus which may be suitable source of antigens is the Hepatitis delta virus. Still other viral sources may include avian infectious bursal disease virus and porcine respiratory and reproductive syndrome virus. The alphavirus family includes equine arteritis virus and various Encephalitis viruses.

[0114] The viruses of the present invention may also carry immunogens which are useful to immunize a human or non-human animal against other pathogens including bacteria, fungi, parasitic microorganisms or multicellular parasites which infect human and non-human vertebrates, or from a cancer cell or tumor cell. Examples of bacterial pathogens include pathogenic gram-positive cocci include pneumococci; staphylococci; and streptococci. Pathogenic gram-negative cocci include meningococcus; gonococcus. Pathogenic enteric gram-negative bacilli include enterobacteriaceae; *pseudomonas*, *acinetobacteria* and *eikenella*; *melioidosis*; *salmonella*; *shigella*; *haemophilus*; *moraxella*; *H. ducreyi* (which causes chancroid); *brucella*; *Francisella tularensis* (which causes tularemia); *yersinia* (pasteurel la); streptobacillus moniliformis and spirillum; Gram-positive bacilli include *listeria monocytogenes*; *erysipelothrix rhusiopathiae*; *Corynebacterium diphtheria* (diphtheria); cholera; *B. anthracis* (anthrax); donovanosis (granuloma inguinale); and bartonellosis. Diseases caused by pathogenic anaerobic bacteria include tetanus; botulism; other clostridia; tuberculosis; leprosy; and other mycobacteria. Pathogenic spirochetal diseases include syphilis; trepone-

matoses; yaws, pinta and endemic syphilis; and leptospirosis. Other infections caused by higher pathogen bacteria and pathogenic fungi include actinomycosis; nocardiosis; cryptococcosis, blastomycosis, histoplasmosis and coccidioidomycosis; candidiasis, aspergillosis, and mucormycosis; sporotrichosis; paracoccidioidomycosis, petriellidiosis, torulopsosis, mycetoma and chromomycosis; and dermatophytosis. Rickettsial infections include Typhus fever, Rocky Mountain spotted fever, Q fever, and Rickettsialpox. Examples of mycoplasma and chlamydial infections include: mycoplasma pneumoniae; lymphogranuloma venereum; psittacosis; and perinatal chlamydial infections. Pathogenic eukaryotes encompass pathogenic protozoans and helminths and infections produced thereby include: amebiasis; malaria; leishmaniasis; trypanosomiasis; toxoplasmosis; *Pneumocystis carinii*; Trichans; *Toxoplasma gondii*; babesiosis; giardiasis; trichinosis; filariasis; schistosomiasis; nematodes; trematodes or flukes; and cestode (tapeworm) infections.

[0115] Many of these organisms and/or toxins produced thereby have been identified by the Centers for Disease Control [(CDC), Department of Health and Human Services, USA], as agents which have potential for use in biological attacks. For example, some of these biological agents, include, *Bacillus anthracis* (anthrax), *Clostridium botulinum* and its toxin (botulism), *Yersinia pestis* (plague), variola major (smallpox), *Francisella tularensis* (tularemia), and viral hemorrhagic fevers [filoviruses (e.g., Ebola, Marburg), and arenaviruses [e.g., Lassa, Machupo]], all of which are currently classified as Category A agents; *Coxiella burnetii* (Q fever); *Brucella* species (brucellosis), *Burkholderia mallei* (glanders), *Burkholderia pseudomallei* (melioidosis), *Ricinus communis* and its toxin (ricin toxin), *Clostridium perfringens* and its toxin (epsilon toxin), *Staphylococcus* species and their toxins (enterotoxin B), *Chlamydia psittaci* (psittacosis), water safety threats (e.g., *Vibrio cholerae*, *Cryptosporidium parvum*), Typhus fever (*Rickettsia powazekii*), and viral encephalitis (alphaviruses, e.g., Venezuelan equine encephalitis; eastern equine encephalitis; western equine encephalitis); all of which are currently classified as Category B agents; and Nipaviruses and hantaviruses, which are currently classified as Category C agents. In addition, other organisms, which are so classified or differently classified, may be identified and/or used for such a purpose in the future. It will be readily understood that the viral vectors and other constructs described herein are useful to deliver antigens from these organisms, viruses, their toxins or other by-products, which will prevent and/or treat infection or other adverse reactions with these biological agents.

[0116] Administration of the vectors of the invention to deliver immunogens against the variable region of the T cells elicit an immune response including CTLs to eliminate those T cells. In RA, several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-3, V-14, V-17 and V $\alpha$ -17. Thus, delivery of a nucleic acid sequence that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in RA. In MS, several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-7 and V $\alpha$ -10. Thus, delivery of a nucleic acid sequence that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in MS. In

scleroderma, several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-6, V-8, V-14 and V $\alpha$ -16, V $\alpha$ -3C, V $\alpha$ -7, V $\alpha$ -14, V $\alpha$ -15, V $\alpha$ -16, V $\alpha$ -28 and V $\alpha$ -12. Thus, delivery of a chimeric adenovirus that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in scleroderma.

**[0117]** C. Ad-Mediated Delivery Methods

**[0118]** The therapeutic levels, or levels of immunity, of the selected gene can be monitored to determine the need, if any, for boosters. Following an assessment of CD8+ T cell response, or optionally, antibody titers, in the serum, optional booster immunizations may be desired. Optionally, the adenoviral vectors of the invention may be delivered in a single administration or in various combination regimens, e.g., in combination with a regimen or course of treatment involving other active ingredients or in a prime-boost regimen. A variety of such regimens have been described in the art and may be readily selected.

**[0119]** For example, prime-boost regimens may involve the administration of a DNA (e.g., plasmid) based vector to prime the immune system to a second or further, booster, administration with a traditional antigen, such as a protein or a recombinant virus carrying the sequences encoding such an antigen. See, e.g., WO 00/11140, published Mar. 2, 2000, incorporated by reference. Alternatively, an immunization regimen may involve the administration of a chimeric adenoviral vector of the invention to boost the immune response to a vector (either viral or DNA-based) carrying an antigen, or a protein. In still another alternative, an immunization regimen involves administration of a protein followed by booster with a vector encoding the antigen.

**[0120]** In one embodiment, the invention provides a method of priming and boosting an immune response to a selected antigen by delivering a plasmid DNA vector carrying said antigen, followed by boosting with an adenoviral vector of the invention. In one embodiment, the prime-boost regimen involves the expression of multiproteins from the prime and/or the boost vehicle. See, e.g., R. R. Amara, *Science*, 292:69-74 (6 April 2001) which describes a multiprotein regimen for expression of protein subunits useful for generating an immune response against HIV and SIV. For example, a DNA prime may deliver the Gag, Pol, Vif, VPX and Vpr and Env, Tat, and Rev from a single transcript. Alternatively, the SIV Gag, Pol and HIV-1 Env is delivered in a recombinant adenovirus construct of the invention. Still other regimens are described in WO 99/16884 and WO 01/54719.

**[0121]** However, the prime-boost regimens are not limited to immunization for HIV or to delivery of these antigens. For example, priming may involve delivering with a first vector of the invention followed by boosting with a second vector, or with a composition containing the antigen itself in protein form. In one example, the prime-boost regimen can provide a protective immune response to the virus, bacteria or other organism from which the antigen is derived. In another desired embodiment, the prime-boost regimen provides a therapeutic effect that can be measured using convention assays for detection of the presence of the condition for which therapy is being administered.

**[0122]** The priming composition may be administered at various sites in the body in a dose dependent manner, which

depends on the antigen to which the desired immune response is being targeted. The invention is not limited to the amount or situs of injection(s) or to the pharmaceutical carrier. Rather, the regimen may involve a priming and/or boosting step, each of which may include a single dose or dosage that is administered hourly, daily, weekly or monthly, or yearly. As an example, the mammals may receive one or two doses containing between about 10  $\mu$ g to about 50  $\mu$ g of plasmid in carrier. A desirable amount of a DNA composition ranges between about 1  $\mu$ g to about 10,000  $\mu$ g of the DNA vector. Dosages may vary from about 1  $\mu$ g to 1000  $\mu$ g DNA per kg of subject body weight. The amount or site of delivery is desirably selected based upon the identity and condition of the mammal.

**[0123]** The dosage unit of the vector suitable for delivery of the antigen to the mammal is described herein. The vector is prepared for administration by being suspended or dissolved in a pharmaceutically or physiologically acceptable carrier such as isotonic saline; isotonic salts solution or other formulations that will be apparent to those skilled in such administration. The appropriate carrier will be evident to those skilled in the art and will depend in large part upon the route of administration. The compositions of the invention may be administered to a mammal according to the routes described above, in a sustained release formulation using a biodegradable biocompatible polymer, or by on-site delivery using micelles, gels and liposomes. Optionally, the priming step of this invention also includes administering with the priming composition, a suitable amount of an adjuvant, such as are defined herein.

**[0124]** Preferably, a boosting composition is administered about 2 to about 27 weeks after administering the priming composition to the mammalian subject. The administration of the boosting composition is accomplished using an effective amount of a boosting composition containing or capable of delivering the same antigen as administered by the priming DNA vaccine. The boosting composition may be composed of a recombinant viral vector derived from the same viral source (e.g., adenoviral sequences of the invention) or from another source. Alternatively, the "boosting composition" can be a composition containing the same antigen as encoded in the priming DNA vaccine, but in the form of a protein or peptide, which composition induces an immune response in the host. In another embodiment, the boosting composition contains a DNA sequence encoding the antigen under the control of a regulatory sequence directing its expression in a mammalian cell, e.g., vectors such as well-known bacterial or viral vectors. The primary requirements of the boosting composition are that the antigen of the composition is the same antigen, or a cross-reactive antigen, as that encoded by the priming composition.

**[0125]** In another embodiment, the adenoviral vectors of the invention are also well suited for use in a variety of other immunization and therapeutic regimens. Such regimens may involve delivery of adenoviral vectors of the invention simultaneously or sequentially with Ad vectors of different serotype capsids, regimens in which adenoviral vectors of the invention are delivered simultaneously or sequentially with non-Ad vectors, regimens in which the adenoviral vectors of the invention are delivered simultaneously or sequentially with proteins, peptides, and/or other biologically useful therapeutic or immunogenic compounds. Such uses will be readily apparent to one of skill in the art.

## IV. Simian Adenovirus 18 Sequences

[0126] The invention provides nucleic acid sequences and amino acid sequences of Ad SA1 8, which are isolated from the other viral material with which they are associated in nature. These sequences are useful in preparing heterologous molecules containing the nucleic acid sequences and amino acid sequences, and regions or fragments thereof as are described herein, viral vectors which are useful for a variety of purposes, including the constructs and compositions, and such methods as are described herein for the chimeric adenoviruses, including, e.g., in host cells for production of viruses requiring adenoviral helper functions, as delivery vehicles for heterologous molecules such as those described herein. These sequences are also useful in generating the chimeric adenoviruses of the invention.

## [0127] A. Nucleic Acid Sequences

[0128] The SA18 nucleic acid sequences of the invention include nucleotides SEQ ID NO: 12, nt 1 to 31967. See, Sequence Listing, which is incorporated by reference herein. The nucleic acid sequences of the invention further encompass the strand which is complementary to the sequences of SEQ ID NO: 12, as well as the RNA and cDNA sequences corresponding to the sequences of these sequences figures and their complementary strands. Further included in this invention are nucleic acid sequences which are greater than 95 to 98%, and more preferably about 99 to 99.9% homologous or identical to the Sequence Listing. Also included in the nucleic acid sequences of the invention are natural variants and engineered modifications of the sequences provided in SEQ ID NO: 12 and their complementary strands. Such modifications include, for example, labels that are known in the art, methylation, and substitution of one or more of the naturally occurring nucleotides with a degenerate nucleotide.

[0129] The invention further encompasses fragments of the sequences of SA 18, their complementary strand, cDNA and RNA complementary thereto. Suitable fragments are at least 15 nucleotides in length, and encompass functional fragments, i.e., fragments which are of biological interest. For example, a functional fragment can express a desired adenoviral product or may be useful in production of recombinant viral vectors. Such fragments include the gene sequences and fragments listed in the tables below.

[0130] The following tables provide the transcript regions and open reading frames in the simian adenovirus sequences of the invention. For certain genes, the transcripts and open reading frames (ORFs) are located on the strand complementary to that presented in SEQ ID NO: 12. See, e.g., E2b, E4 and E2a. The calculated molecular weights of the encoded proteins are also shown.

| Adenovirus Gene | Protein | Ad SA18, SEQ ID NO: 12 |      |       |
|-----------------|---------|------------------------|------|-------|
|                 |         | start                  | End  | M.W.  |
| ITR             |         | 1                      | 180  |       |
| E1a             | 13S     | 916                    | 1765 | 27264 |
|                 | 12S     | 916                    | 1765 | 24081 |

-continued

| Adenovirus Gene | Protein      | Ad SA18, SEQ ID NO: 12 |       |        |
|-----------------|--------------|------------------------|-------|--------|
|                 |              | start                  | End   | M.W.   |
| E1b             | Small T      | 1874                   | 2380  | 19423  |
|                 | LargeT       | 2179                   | 3609  | 52741  |
|                 | IX           | 3678                   | 4079  | 13701  |
| E2b             | IVa2         | 5478                   | 4126  | 51295  |
|                 | Polymerase   | 13745                  | 5229  | 128392 |
|                 | PTP          | 13745                  | 8597  | 75358  |
| L1              | Agnoprotein  | 8007                   | 8705  | 23610  |
|                 | 52/55 kD     | 10788                  | 11945 | 43416  |
|                 | IIIa         | 11966                  | 13699 | 63999  |
| L2              | Penton       | 13796                  | 15322 | 57166  |
|                 | VII          | 15328                  | 15873 | 20352  |
| L3              | V            | 15920                  | 17050 | 42020  |
|                 | VI           | 17348                  | 18154 | 29222  |
|                 | Hexon        | 18257                  | 21010 | 102912 |
|                 | Endoprotease | 21029                  | 21640 | 23015  |

## [0131]

| Adenovirus Gene Region | Protein       | Ad SA18, SEQ ID NO: 12 |       |       |
|------------------------|---------------|------------------------|-------|-------|
|                        |               | start                  | End   | M.W.  |
| 2a                     | DBP           | 23147                  | 21711 | 53626 |
| L4                     | 100 kD        | 23175                  | 25541 | 87538 |
|                        | 22 kD homolog | 25204                  | 25797 | 22206 |
|                        | 33 kD homolog | 25204                  | 26025 | 24263 |
|                        | VIII          | 26107                  | 26817 | 25490 |
| E3                     | Orf #1        | 26817                  | 27125 | 11814 |
| L5                     | Fiber         | 27192                  | 29015 | 65455 |
| E4                     | Orf 6/7       | 30169                  | 29067 | 13768 |
|                        | Orf 6         | 30169                  | 29303 | 33832 |
|                        | Orf 4         | 30464                  | 30099 | 14154 |
|                        | Orf 3         | 30816                  | 30466 | 13493 |
|                        | Orf 2         | 31205                  | 30813 | 14698 |
|                        | Orf 1         | 31608                  | 31231 | 14054 |
| ITR                    |               | 31788                  | 31967 |       |

[0132] The SA18 adenoviral nucleic acid sequences are useful as therapeutic and immunogenic agents and in construction of a variety of vector systems and host cells. Such vectors are useful for any of the purposes described above for the chimeric adenovirus. Additionally, these SA18 sequences and products may be used alone or in combination with other adenoviral sequences or fragments, or in combination with elements from other adenoviral or non-adenoviral sequences. The adenoviral sequences of the invention are also useful as antisense delivery vectors, gene therapy vectors, or vaccine vectors, and in methods of using same. Thus, the invention further provides nucleic acid molecules, gene delivery vectors, and host cells which contain the Ad sequences of the invention.

[0133] For example, the invention encompasses a nucleic acid molecule containing simian Ad ITR sequences of the invention. In another example, the invention provides a nucleic acid molecule containing simian Ad sequences of the invention encoding a desired Ad gene product. Still other nucleic acid molecule constructed using the sequences of the invention will be readily apparent to one of skill in the art, in view of the information provided herein.

[0134] In one embodiment, the simian Ad gene regions identified herein may be used in a variety of vectors for delivery of a heterologous molecule to a cell. Examples of such molecules and methods of delivery are provided in Section III herein. For example, vectors are generated for expression of an adenoviral capsid protein (or fragment thereof) for purposes of generating a viral vector in a packaging host cell. Such vectors may be designed for expression in trans. Alternatively, such vectors are designed to provide cells which stably contain sequences which express desired adenoviral functions, e.g., one or more of E1a, E1b, the terminal repeat sequences, E2a, E2b, E4, E4ORF6 region.

[0135] In addition, the adenoviral gene sequences and fragments thereof are useful for providing the helper functions necessary for production of helper-dependent viruses (e.g., adenoviral vectors deleted of essential functions or adeno-associated viruses (AAV)). For such production methods, the simian adenoviral sequences of the invention are utilized in such a method in a manner similar to those described for the human Ad. However, due to the differences in sequences between the simian adenoviral sequences of the invention and those of human Ad, the use of the sequences of the invention essentially eliminate the possibility of homologous recombination with helper functions in a host cell carrying human Ad E1 functions, e.g., 293 cells, which may produce infectious adenoviral contaminants during rAAV production.

[0136] Methods of producing rAAV using adenoviral helper functions have been described at length in the literature with human adenoviral serotypes. See, e.g., U.S. Pat. No. 6,258,595 and the references cited therein. See, also, U.S. Pat. No. 5,871,982; WO 99/14354; WO 99/15685; WO 99/47691. These methods may also be used in production of non-human serotype AAV, including non-human primate AAV serotypes. The simian adenoviral gene sequences of the invention which provide the necessary helper functions (e.g., E1a, E1b, E2a and/or E4 ORF6) can be particularly useful in providing the necessary adenoviral function while minimizing or eliminating the possibility of recombination with any other adenoviruses present in the rAAV-packaging cell which are typically of human origin. Thus, selected genes or open reading frames of the adenoviral sequences of the invention may be utilized in these rAAV production methods.

[0137] Alternatively, recombinant adenoviral simian vectors of the invention may be utilized in these methods. Such recombinant adenoviral simian vectors may include, e.g., a hybrid simian Ad/AAV in which simian Ad sequences flank a rAAV expression cassette composed of, e.g., AAV 3' and/or 5' ITRs and a transgene under the control of regulatory sequences which control its expression. One of skill in the art will recognize that still other simian adenoviral vectors and/or gene sequences of the invention will be useful for production of rAAV and other viruses dependent upon adenoviral helper.

[0138] In still another embodiment, nucleic acid molecules are designed for delivery and expression of selected adenoviral gene products in a host cell to achieve a desired physiologic effect. For example, a nucleic acid molecule containing sequences encoding an adenovirus E1a protein of the invention may be delivered to a subject for use as a

cancer therapeutic. Optionally, such a molecule is formulated in a lipid-based carrier and preferentially targets cancer cells. Such a formulation may be combined with other cancer therapeutics (e.g., cisplatin, taxol, or the like). Still other uses for the adenoviral sequences provided herein will be readily apparent to one of skill in the art.

[0139] In addition, one of skill in the art will readily understand that the Ad sequences of the invention can be readily adapted for use for a variety of viral and non-viral vector systems for in vitro, ex vivo or in vivo delivery of therapeutic and immunogenic molecules, including any of those identified as being deliverable via the chimeric adenoviruses of the invention. For example, the simian Ad genome of the invention can be utilized in a variety of rAd and non-rAd vector systems. Such vector systems may include, e.g., plasmids, lentiviruses, retroviruses, poxviruses, vaccinia viruses, and adeno-associated viral systems, among others. Selection of these vector systems is not a limitation of the present invention.

[0140] The invention further provides molecules useful for production of the simian and simian-derived proteins of the invention. Such molecules which carry polynucleotides including the simian Ad DNA sequences of the invention can be in the form of a vector.

#### [0141] B. Simian Adenoviral Proteins of the Invention

[0142] The invention further provides gene products of the above adenoviruses, such as proteins, enzymes, and fragments thereof, which are encoded by the adenoviral nucleic acids of the invention. The invention further encompasses SA18 proteins, enzymes, and fragments thereof, having the amino acid sequences encoded by these nucleic acid sequences which are generated by other methods. Such proteins include those encoded by the open reading frames identified in the tables above, and fragments thereof.

[0143] Thus, in one aspect, the invention provides unique simian adenoviral proteins which are substantially pure, i.e., are free of other viral and proteinaceous proteins. Preferably, these proteins are at least 10% homogeneous, more preferably 60% homogeneous, and most preferably 95% homogeneous.

[0144] In one embodiment, the invention provides unique simian-derived capsid proteins. As used herein, a simian-derived capsid protein includes any adenoviral capsid protein that contains a SA18 capsid protein or a fragment thereof, as defined above, including, without limitation, chimeric capsid proteins, fusion proteins, artificial capsid proteins, synthetic capsid proteins, and recombinantly capsid proteins, without limitation to means of generating these proteins.

[0145] Suitably, these simian-derived capsid proteins contain one or more SA18 regions or fragments thereof (e.g., a hexon, penton, fiber or fragment thereof) in combination with capsid regions or fragments thereof of different adenoviral serotypes, or modified simian capsid proteins or fragments, as described herein. A "modification of a capsid protein associated with altered tropism" as used herein includes an altered capsid protein, i.e., a penton, hexon or fiber protein region, or fragment thereof, such as the knob domain of the fiber region, or a polynucleotide encoding same, such that specificity is altered. The simian-derived capsid may be constructed with one or more of the simian



Ad of the invention or another Ad serotypes which may be of human or non-human origin. Such Ad may be obtained from a variety of sources including the ATCC, commercial and academic sources, or the sequences of the Ad may be obtained from GenBank or other suitable sources.

[0146] The amino acid sequences of the simian adenoviruses penton proteins of the invention are provided herein. The AdSA18 penton protein is provided in SEQ ID NO: 13. Suitably, any of these penton proteins, or unique fragments thereof, may be utilized for a variety of purposes. Examples of suitable fragments include the penton having N-terminal and/or C-terminal truncations of about 50, 100, 150, or 200 amino acids, based upon the amino acid numbering provided above. Other suitable fragments include shorter internal, C-terminal, or N-terminal fragments. Further, the penton protein may be modified for a variety of purposes known to those of skill in the art.

[0147] The invention further provides the amino acid sequences of the hexon protein of SA18, SEQ ID NO:14. Suitably, this hexon protein, or unique fragments thereof, may be utilized for a variety of purposes. Examples of suitable fragments include the hexon having N-terminal and/or C-terminal truncations of about 50, 100, 150, 200, 300, 400, or 500 amino acids, based upon the amino acid numbering provided above and in SEQ ID NO: 14. Other suitable fragments include shorter internal, C-terminal, or N-terminal fragments. For example, one suitable fragment the loop region (domain) of the hexon protein, designated DE1 and FG1, or a hypervariable region thereof. Such fragments include the regions spanning amino acid residues about 125 to 443; about 138 to 441, or smaller fragments, such as those spanning about residue 138 to residue 163; about 170 to about 176; about 195 to about 203; about 233 to about 246; about 253 to about 264; about 287 to about 297; about 404 to about 430, about 430 to 550, about 545 to 650; of the simian hexon proteins, with reference to SEQ ID NO: 14. Other suitable fragments may be readily identified by one of skill in the art. Further, the hexon protein may be modified for a variety of purposes known to those of skill in the art. Because the hexon protein is the determinant for serotype of an adenovirus, such artificial hexon proteins would result in adenoviruses having artificial serotypes. Other artificial capsid proteins can also be constructed using the chimp Ad penton sequences and/or fiber sequences of the invention and/or fragments thereof.

[0148] In one example, it may be desirable to generate an adenovirus having an altered hexon protein utilizing the sequences of a hexon protein of the invention. One suitable method for altering hexon proteins is described in U.S. Pat. No. 5,922,315, which is incorporated by reference. In this method, at least one loop region of the adenovirus hexon is changed with at least one loop region of another adenovirus serotype. Thus, at least one loop region of such an altered adenovirus hexon protein is a simian Ad hexon loop region of the invention. In one embodiment, a loop region of the SA18 hexon protein is replaced by a loop region from another adenovirus serotype. In another embodiment, the loop region of the SA18 hexon is used to replace a loop region from another adenovirus serotype. Suitable adenovirus serotypes may be readily selected from among human and non-human serotypes, as described herein. SA18 is selected for purposes of illustration only; the other simian Ad hexon proteins of the invention may be similarly altered,

or used to alter another Ad hexon. The selection of a suitable serotype is not a limitation of the present invention. Still other uses for the hexon protein sequences of the invention will be readily apparent to those of skill in the art.

[0149] The invention further encompasses the fiber proteins of the simian adenoviruses of the invention. The fiber protein of AdSA18 has the amino acid sequence of SEQ ID NO: 15. Suitably, this fiber protein, or unique fragments thereof, may be utilized for a variety of purposes. One suitable fragment is the fiber knob, which spans about amino acids 247 to 425 of SEQ ID NO: 15. Examples of other suitable fragments include the fiber having N-terminal and/or C-terminal truncations of about 50, 100, 150, or 200 amino acids, based upon the amino acid numbering provided above and in SEQ ID NO: 15. Still other suitable fragments include internal fragments. Further, the fiber protein may be modified using a variety of techniques known to those of skill in the art.

[0150] The invention further encompasses unique fragments of the proteins of the invention which are at least 8 amino acids in length. However, fragments of other desired lengths can be readily utilized. In addition, the invention encompasses such modifications as may be introduced to enhance yield and/or expression of an SA18 gene product, e.g., construction of a fusion molecule in which all or a fragment of the SA18 gene product is fused (either directly or via a linker) with a fusion partner to enhance. Other suitable modifications include, without limitation, truncation of a coding region (e.g., a protein or enzyme) to eliminate a pre- or pro-protein ordinarily cleaved and to provide the mature protein or enzyme and/or mutation of a coding region to provide a secretable gene product. Still other modifications will be readily apparent to one of skill in the art. The invention further encompasses proteins having at least about 95% to 99% identity to the SA 18 proteins provided herein.

[0151] As described herein, vectors of the invention containing the adenoviral capsid proteins of the invention are particularly well suited for use in applications in which the neutralizing antibodies diminish the effectiveness of other Ad serotype based vectors, as well as other viral vectors. The rAd vectors of the invention are particularly advantageous in readministration for repeat gene therapy or for boosting immune response (vaccine titers). Examples of such regimens are provided herein.

[0152] Under certain circumstances, it may be desirable to use one or more of the SA18 gene products (e.g., a capsid protein or a fragment thereof) to generate an antibody. The term "an antibody," as used herein, refers to an immunoglobulin molecule which is able to specifically bind to an epitope. Thus, the antibodies of the invention bind, preferably specifically and without cross-reactivity, to a SA18 epitope. The antibodies in the present invention exist in a variety of forms including, for example, high affinity polyclonal antibodies, monoclonal antibodies, synthetic antibodies, chimeric antibodies, recombinant antibodies and humanized antibodies. Such antibodies originate from immunoglobulin classes IgG, IgM, IgA, IgD and IgE.

[0153] Such antibodies may be generated using any of a number of methods known in the art. Suitable antibodies may be generated by well-known conventional techniques, e.g. Kohler and Milstein and the many known modifications

thereof. Similarly desirable high titer antibodies are generated by applying known recombinant techniques to the monoclonal or polyclonal antibodies developed to these antigens [see, e.g., PCT Patent Application No. PCT/GB85/00392; British Patent Application Publication No. GB2188638A; Amit et al., 1986 *Science*, 233:747-753; Queen et al., 1989 *Proc. Nat'l. Acad. Sci. USA*, 86:10029-10033; PCT Patent Application No. PCT/WO9007861; and Riechmann et al., *Nature*, 332:323-327 (1988); Huse et al., 1988a *Science*, 246:1275-1281]. Alternatively, antibodies can be produced by manipulating the complementarity determining regions of animal or human antibodies to the antigen of this invention. See, e.g., E. Mark and Padlin, "Humanization of Monoclonal Antibodies", Chapter 4, *The Handbook of Experimental Pharmacology*, Vol. 113, *The Pharmacology of Monoclonal Antibodies*, Springer-Verlag (June, 1994); Harlow et al., 1999, *Using Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, N.Y.; Harlow et al., 1989, *Antibodies: A Laboratory Manual*, Cold Spring Harbor, N.Y.; Houston et al., 1988, *Proc. Natl. Acad. Sci. USA* 85:5879-5883; and Bird et al., 1988, *Science* 242:423-426. Further provided by the present invention are anti-idiotypic antibodies (Ab2) and anti-anti-idiotypic antibodies (Ab3). See, e.g., M. Wettendorff et al., "Modulation of anti-tumor immunity by anti-idiotypic antibodies." In *Idiotypic Network and Diseases*, ed. by J. Cerny and J. Hiernaux, 1990 *J. Am. Soc. Microbiol.*, Washington D.C.: pp. 203-229]. These anti-idiotypic and anti-anti-idiotypic antibodies are produced using techniques well known to those of skill in the art. These antibodies may be used for a variety of purposes, including diagnostic and clinical methods and kits.

[0154] Under certain circumstances, it may be desirable to introduce a detectable label or a tag onto a SA18 gene product, antibody or other construct of the invention. As used herein, a detectable label is a molecule which is capable, alone or upon interaction with another molecule, of providing a detectable signal. Most desirably, the label is detectable visually, e.g. by fluorescence, for ready use in immunohistochemical analyses or immunofluorescent microscopy. For example, suitable labels include fluorescein isothiocyanate (FITC), phycoerythrin (PE), allophycocyanin (APC), coriophosphine-O (CPO) or tandem dyes, PE-cyanin-5 (PC5), and PE-Texas Red (ECD). All of these fluorescent dyes are commercially available, and their uses known to the art. Other useful labels include a colloidal gold label. Still other useful labels include radioactive compounds or elements. Additionally, labels include a variety of enzyme systems that operate to reveal a calorimetric signal in an assay, e.g., glucose oxidase (which uses glucose as a substrate) releases peroxide as a product which in the presence of peroxidase and a hydrogen donor such as tetramethyl benzidine (TMB) produces an oxidized TMB that is seen as a blue color. Other examples include horseradish peroxidase (HRP) or alkaline phosphatase (AP), and hexokinase in conjunction with glucose-6-phosphate dehydrogenase which reacts with ATP, glucose, and NAD<sup>+</sup> to yield, among other products, NADH that is detected as increased absorbance at 340 nm wavelength.

[0155] Other label systems that are utilized in the methods of this invention are detectable by other means, e.g., colored latex microparticles [Bangs Laboratories, Indiana] in which a dye is embedded are used in place of enzymes to form

conjugates with the target sequences provide a visual signal indicative of the presence of the resulting complex in applicable assays.

[0156] Methods for coupling or associating the label with a desired molecule are similarly conventional and known to those of skill in the art. Known methods of label attachment are described [see, for example, *Handbook of Fluorescent probes and Research Chemicals*, 6th Ed., R. P. M. Haugland, Molecular Probes, Inc., Eugene, Oreg., 1996; *Pierce Catalog and Handbook, Life Science and Analytical Research Products*, Pierce Chemical Company, Rockford, Ill., 1994/1995]. Thus, selection of the label and coupling methods do not limit this invention.

[0157] The sequences, proteins, and fragments of the invention may be produced by any suitable means, including recombinant production, chemical synthesis, or other synthetic means. Suitable production techniques are well known to those of skill in the art. See, e.g., Sambrook et al, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press (Cold Spring Harbor, N.Y.). Alternatively, peptides can also be synthesized by the well known solid phase peptide synthesis methods (Merrifield, *J. Am. Chem. Soc.*, 85:2149 (1962); Stewart and Young, *Solid Phase Peptide Synthesis* (Freeman, San Francisco, 1969) pp. 27-62). These and other suitable production methods are within the knowledge of those of skill in the art and are not a limitation of the present invention.

[0158] In addition, one of skill in the art will readily understand that the Ad sequences of the invention can be readily adapted for use for a variety of viral and non-viral vector systems for in vitro, ex vivo or in vivo delivery of therapeutic and immunogenic molecules. For example, in one embodiment, the simian Ad capsid proteins and other simian adenovirus proteins described herein are used for non-viral, protein-based delivery of genes, proteins, and other desirable diagnostic, therapeutic and immunogenic molecules. In one such embodiment, a protein of the invention is linked, directly or indirectly, to a molecule for targeting to cells with a receptor for adenoviruses. Preferably, a capsid protein such as a hexon, penton, fiber or a fragment thereof having a ligand for a cell surface receptor is selected for such targeting. Suitable molecules for delivery are selected from among the therapeutic molecules described herein and their gene products. A variety of linkers including, lipids, polyLys, and the like may be utilized as linkers. For example, the simian penton protein may be readily utilized for such a purpose by production of a fusion protein using the simian penton sequences in a manner analogous to that described in Medina-Kauwe L K, et al, *Gene Ther.* 2001 May; 8(10):795-803 and Medina-Kauwe L K, et al, *Gene Ther.* 2001 Dec; 8(23): 1753-1761. Alternatively, the amino acid sequences of simian Ad protein IX may be utilized for targeting vectors to a cell surface receptor, as described in US Patent Appln 20010047081. Suitable ligands include a CD40 antigen, an RGD-containing or polylysine-containing sequence, and the like. Still other simian Ad proteins, including, e.g., the hexon protein and/or the fiber protein, may be used for used for these and similar purposes.

[0159] Still other adenoviral proteins of the invention may be used as alone, or in combination with other adenoviral protein, for a variety of purposes which will be readily

apparent to one of skill in the art. In addition, still other uses for the adenoviral proteins of the invention will be readily apparent to one of skill in the art.

[0160] The compositions of this invention include vectors that deliver a heterologous molecule to cells, either for therapeutic or vaccine purposes. Such vectors, containing simian adenovirus DNA of SA18 and a minigene, can be constructed using techniques such as those described herein for the chimeric adenoviruses and such techniques as are known in the art. Alternatively, SA19 may be a source for sequences of the chimeric adenoviruses are described herein.

[0161] The following examples are illustrative, and are not intended to limit the invention to those illustrated embodiments.

#### EXAMPLE 1

##### Construction of Pan5/C1 Chimeric Simian Viruses

[0162] Five different adenoviruses initially isolated from the chimpanzee, AdC68 [U.S. Pat. No. 6,083,716], AdPan5, AdPan7, AdPan6 and AdC1 [U.S. Pat. No. 6,083,716] have been sequenced. See, International Application No. PCT/US02/33645, filed November 2002 for the sequences of Pan5 [SEQ ID NO:1], Pan7 [SEQ ID NO:3], and Pan6 [SEQ ID NO:2]. This application also provides sequences for SV1, SV25 and SV39 [SEQ ID No. 4, 5, 6, respectively]. Sequence comparison of the capsid protein sequences predicted that AdC1 clearly belonged to a different serological subgroup than the other four chimpanzee derived adenoviruses.

[0163] However, attempts to cultivate AdC1 in HEK293 cells revealed it to be fastidious in its growth characteristics (data not shown) and therefore possibly unsuitable for use as a vector using the currently available E1 complementing cell lines. However, because of the obvious sequence dissimilarity of AdC1 capsid protein sequence from the other chimpanzee derived adenoviruses (as well as the huAd5), chimeric adenovirus vectors were generated with the capsid characteristics of AdC1. In view of the above-mentioned drawbacks associated with only making hexon changes, more extensive replacements were made in the chimera described herein, i.e., construction of chimeras where the replacement went beyond just the hexon, to achieve two goals. The first was to determine whether making extended replacements would allow for the rescue of viruses containing hexons of unrelated serotypes that may not otherwise be amenable to rescue. The second goal was to test whether the growth characteristics of adenovirus vectors such as AdPan5, that have been found in our laboratory to be able to be grown to high titer for the purpose of manufacture, would also be present in the chimeric virus, particularly when the hexon (and other capsid proteins) are derived from a virus such as AdC1 that are difficult to grow to a high yield in cell lines such as HEK293. An added bonus of extending the replacement to include the fiber protein would be to further increase the antigenic dissimilarity to beyond that afforded by a hexon change alone.

[0164] A. Construction of Two Pan5/C1 Chimeric Plasmids

[0165] The overall approach towards constructing chimeric viruses was to first assemble the complete E1 deleted virus DNA into a single plasmid flanked by recognition sites for the restriction enzyme *Swa*I, digest the plasmid DNA with *Swa*I to release the virus DNA ends, and transfect the DNA into HEK293 cells to determine whether viable chimeric adenovirus could be rescued. Two chimeric virus plasmids were constructed, p5C1short and p5C1long.

[0166] The plasmid p5C1short harbors an E1 deleted Pan5 virus where an internal 15226 bp segment (18332-33557) has been replaced by a functionally analogous 14127 bp (18531-32657) from AdC1. This results in the replacement of the Pan5 proteins hexon, endoprotease, DNA binding protein, 100 kD scaffolding protein, 33 kD protein, protein VIII, and fiber, as well as the entire E3 region, with the homologous segment from AdC1. The C1aI site at the left end of the AdC1 fragment is at the beginning of the hexon gene and the resulting protein is identical to the C1 hexon. The *Eco*RI site which constitutes the right end of the AdC1 fragment is within the E4 orf 7 part of the AdC1. The right end was ligated to a PCR generated right end fragment from AdPan5 such that the regenerated orf 7-translation product is chimeric between AdPan5 and AdC1.

[0167] The plasmid p5C1long harbors an E1 deleted Pan5 virus where an internal 25603 bp segment (7955-33557) has been replaced by a functionally analogous 24712 bp (7946-32657) from AdC1. This results in the replacement of the AdPan5 pre-terminal protein, 52/55 kD protein, penton base protein, protein VII, Mu, and protein VI with those from AdC1 in addition to those replaced in p5C1short. The *Asc*I site at the left end of the AdC1 fragment is at the beginning of the DNA polymerase gene and results in a chimeric protein where the first 165 amino acids of the AdPan5 DNA polymerase has been replaced by a 167 amino acid segment from AdC1 DNA polymerase. In this N-terminal region, the homology between the AdPan5 and AdC1 DNA polymerase proteins is 81% (72% identity).

[0168] The plasmid pDVP5Mlu which contains the left end of AdPan5 was used as the starting plasmid for the chimeric vector construction.

[0169] The plasmid pDVP5Mlu was made as follows. A synthetic DNA fragment harboring recognition sites for the restriction enzymes *Sma*I, *Mlu*I, *Eco*RI and *Eco*RV respectively was ligated into pBR322 digested with *Eco*RI and *Nde*I so as to retain the origin of replication and the beta-lactamase gene. The left end of Pan5 extending to the *Mlu*I site (15135 bp) was cloned into this plasmid between the *Sma*I and *Mlu*I sites. The E1 gene was functionally deleted and replaced by a DNA fragment harboring recognition sites for the extremely rare cutter restriction enzyme sites I-CeuI and PI-SceI). The 2904 base pairs of the right end of Pan-5 was PCR amplified using the primers P5L [GCG CAC GCG TCT CTA TCG ATG AAT TCC ATT GGT GAT GGA CAT GC, SEQ ID NO:7] and P51TR [GCG CAT TTA AAT CAT CAT CAA TAA TAT ACC TCA AAC, SEQ ID NO:8] using Tgo polymerase (Roche). The PCR product was cut with *Mlu*I and *Swa*I, and cloned between *Mlu*I and *Eco*RV of pDVP5Mlu to yield pPan5Mlu+RE. A 3193 bp fragment extending from the *Mlu*I site (15135) to the *Cla*I (18328) site of Pan5 was then inserted between the same

sites of pPan5Mlu+RE to yield pPan5Cla+RE. The 3671 bp ClaI (18531) to EcoRI (22202) fragment of the adenovirus C1 was cloned into pPan5Cla+RE between ClaI (16111) and EcoRI (16116) to yield pPan5C1delRI. The 10452 bp internal EcoRI fragment of the adenovirus C1 (22202-32653) was cloned into the EcoRI site of pPan5C1delRI to yield p5C1short. To construct p5C1long, the AdC1 replacement was further extended by replacing the AscI-ClaI 10379 bp fragment of AdPan5 in p5C1short with the AdC1 AscI-ClaI 10591 bp fragment. Finally a green fluorescent protein (GFP) expression cassette was inserted into both p5C1short and p5C1long between the I-CeuI and PI-SceI sites to yield p5C1shortGFP and p5C1longGFP respectively.

**[0170]** B. Rescue of Chimeric Pan5/C] Recombinant Vector Adenoviruses

**[0171]** The plasmids p5C1shortGFP and p5C1longGFP were digested with the restriction enzyme SwaI and transfected into HEK 293 cells. A typical adenovirus induced cytopathic effect was observed. The rescue of recombinant chimeric adenovirus from the p5C1longGFP transfection was confirmed by collecting the supernatant from the transfection and re-infecting fresh cells which were found to be transduced as determined by GFP expression. Viral DNA prepared from the chimeric recombinant virus was digested with several restriction enzymes and found to have the expected pattern on electrophoresis (data not shown).

**[0172]** The chimeric adenoviral construct with the shorter replacement p5C1short encodes the C1 proteins hexon and fiber as well as the intervening open reading frames for endoprotease, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, and protein VIII. (The E3 region is also included within this region but is unlikely to impact on the viability of the chimeric virus). When the replacement was extended to include the additional AdC1 proteins pTP (pre-terminal protein), 52/55 kDa protein, penton base, protein VII, Mu, and protein VI, there was no difficulty in rescuing viable chimeric virus. In this experiment, the chimeric adenovirus construction strategy utilized the presence of AscI and ClaI restriction enzyme sites present on the genes for DNA polymerase and hexon respectively on both AdPan5 and AdC1.

**[0173]** The reasons for the relatively higher yield of the chimeric virus compared to the wild-type AdC1 virus are not clear. In the growth of the 5C1 chimeric virus in 293 cells, the adenoviral early region gene products of E1 and E4 are derived from Ad5 and AdPan5 respectively. The E1 and E4 gene products bind, regulate and de-repress several cellular transcription complexes and coordinate their activity towards viral multiplication. Thus it is possible that the E1 gene products supplied in trans from the 293 cells and the E4 gene products from AdPan5 are more optimal in the human 293 cell background than are the equivalent AdC1 gene products. This may also apply to the activity of the major late promoter whose activity is responsible for the transcription of the capsid protein genes. In the chimeric virus, the major late promoter, and the protein IVa2 which transactivates it, are derived from AdPan5. However the E2 gene products required for adenoviral DNA replication pTP and single-stranded DNA—binding protein are derived from AdC1. The adenoviral DNA polymerase, which complexes with pTP, is chimeric in Ad5C1 but mostly AdPan5 derived.

## EXAMPLE 2

### Generation of Simian Pan5/Human Ad40 Chimeric Adenovirus and Chimpanzee Pan5/Simian SA18 Chimeric Adenovirus

**[0174]** The construction of plasmids designed to rescue chimeric adenoviruses where the outside flanking regions are derived from the chimpanzee adenovirus AdPan5, and the internal region (containing the structural capsid protein genes) are derived from the human adenovirus Ad40 and the simian adenovirus SA18, are described below.

**[0175]** As described for the Pan5-C1 chimeric adenovirus, the overall approach towards constructing chimeric viruses was to first assemble the complete E1 deleted virus DNA into a single plasmid flanked by recognition sites for the restriction enzyme SwaI, digest the plasmid DNA with SwaI to release the virus DNA ends, and transfect the DNA into HEK293 cells to determine whether viable chimeric adenovirus could be rescued. Two chimeric virus plasmids were constructed, pPan5-40 and pPan5-SA18 corresponding to the two two chimeric adenoviruses referred to above. The plasmid pPan5-40 harbors an E1 deleted Pan5 virus where an internal 22975 bp segment (10400-33374) has been replaced by a functionally analogous 21603 bp (10043-21603) from Ad40. This results in the replacement of the AdPan5 52/55 kD protein, penton base protein, protein VII, Mu, protein VI, hexon, endoprotease, DNA binding protein, 100 KD scaffolding protein, 33 kD protein, protein VIII, and fiber, as well as the entire E3 region, with the homologous segment from Ad40. Similarly, the plasmid pPan5-SA18 harbors an E1 deleted Pan5 virus where an internal 22975 bp segment (10400-33374) has been replaced by a functionally analogous 19015 bp (10573-29587) from SA18. This results in the replacement of the AdPan5 52/55 kD protein, penton base protein, protein VII, Mu, protein VI, hexon, endoprotease, DNA binding protein, 100 kD scaffolding protein, 33 kD protein, protein VIII, and fiber, as well as the entire E3 region, with the homologous segment from SA18.

**[0176]** The construction of plasmids designed to rescue chimeric adenoviruses where the outside flanking regions are derived from the chimpanzee adenovirus AdPan5, and the internal region (containing the structural capsid protein genes) are derived from the human adenovirus Ad40 and the simian adenovirus SA18, are described below.

**[0177]** A. Silent Mutagenesis of XbaI Site:

**[0178]** The plasmid pDVP5Mlu which contains the left end of AdPan5 was used as the starting plasmid for the chimeric vector construction. As a first step the XbaI site (3820) was mutagenized to destroy the recognition site without changing the coding sequence for polymerase. This was done by first sub-cloning the NdeI (812) to HindIII (4931) fragment into the plasmid pNEB 193 (New England Biolabs) using the same restriction sites in pNEB 193, to yield pNEBp5. A PCR reaction was performed on pNEBp5 using the primers P5XTOP (GATACCTAGGAACGAG-GAGGATTTGATATTG, SEQ ID NO:9) and P5XBOT (ATGTACGCTCCGCGCTCAC, SEQ ID NO:10) to yield a 591 bp product. The PCR product was cleaved with AvrII and BbvCI and ligated into pNEBp5 cut with XbaI and BbvCI to yield the desired mutation in the plasmid pNEBp5mut. The mutated NdeI-HindIII fragment from

pNEBp5mut was ligated back into pDVP5Mlu to yield the desired mutated plasmid pDVP5Mlumut.

[0179] B. Insertion of the Pan5 Right End Comprising the Right ITR and the Complete E4 Region:

[0180] The right end of Pan-5 was PCR amplified (P5RE2PCR) using the primers P5E4 [GATCGAATTC-CCACTCTGTACCCCATCTCTG, SEQ ID NO:11] and P5ITR [GCG CAT TTA AAT CAT CAT CAA TAA TAT ACC TCA AAC, SEQ ID NO:8] using Tgo polymerase, cut with EcoRI and SwaI, and cloned between EcoRI and EcoRV of pDVP5Mlumut to yield pPan5Mlumut+RE.

[0181] C. Insertion of Ad40 or SA18 Structural Protein Sequences:

[0182] In order to construct p5-40, the Ad40 segment from XbaI (10038) to EcoRI (31642) was ligated into pDVP5Mlumut+RE2 between XbaI (8178) and EcoRI (12924) in two steps: first, the XbaI (30494) to EcoRI

(31642) fragment was inserted, followed by the XbaI (10038)-XbaI (30494) fragment. To construct pPan5-SA18 the XbaI (10568) to EcoRI (29584) fragment from the simian adenovirus SA18 was inserted into pDVP5Mlumut+RE2 between XbaI (8178) and EcoRI (12924).

[0183] The minigene encoding for green fluorescent protein was inserted in place of the E1 deletion between the I-CeuI and PI-SceI sites in pPan5-40 and pPan5-SA18 respectively. This plasmids were purified, digested with SwaI and transfected into 293 cells.

[0184] All publications cited in this specification are incorporated herein by reference. While the invention has been described with reference to a particularly preferred embodiment, it will be appreciated that modifications can be made without departing from the spirit of the invention. Such modifications are intended to fall within the scope of the appended claims.

---

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 15

<210> SEQ ID NO 1

<211> LENGTH: 36462

<212> TYPE: DNA

<213> ORGANISM: chimpanzee adenovirus serotype Pan5

<400> SEQUENCE: 1

```

catcatcaat aatatacctc aaacttttgg tgcgcgtaa tatgcaaatg aggtatttga      60
at ttggggat gcggggcggt gattggctgc gggagcggcg accgtaggg gcggggcggg      120
tgacgttttg atgacgtggc cgtgaggcgg agccggtttg caagttctcg tgggaaaagt      180
gacgtcaaac gaggtgtggt ttgaacacgg aaataactcaa ttttcccgcg ctctctgaca      240
ggaaatgagg tgtttctggg cggatgcaag tgaaacggg ccattttcgc gcgaaaactg      300
aatgaggaag tgaaaatctg agtaattccg cgtttatggc agggaggagt atttgccgag      360
ggccgagtag actttgaccg attacgtggg ggtttcgatt accgtat tttt tcaccta aat      420
ttccgcgtac ggtgtca aag tccggtgttt ttacgtagg tgcagctgat cgccagggta      480
tttaaacctg cgctctctag tcaagaggcc actccttgagt gccagcgagt agagt ttttct      540
cctccgcgcc gcgagtcaga tctacacttt gaaagatgag gcacctgaga gacctgcccg      600
gtaatgtttt cctggctact gggaacgaga ttctggaact ggtggtggac gccatgatgg      660
gtgacgaccc tccggagccc cctaccccat ttgaagcggc ttgctgttac gatttgtatg      720
atctggagggt ggatgtgccc gagaacgacc ccaacgagga ggcggtgaat gatttgttta      780
gcgatgccgc gctgctggct gccgagcagg ctaatacggc ctctggctca gacagcgatt      840
cctctctcca taccocgaga cccggcagag gtgagaaaaa gatccccgag cttaaagggg      900
aagagctcga cctgocgtgc tatgaggaat gcttgccctc gagcgatgat gaggagacg      960
aggaggcgat tcgagctgca gcgaaccagg gagtgaaaac agcgagcgag ggctttagcc      1020
tggaactgtc tactctgccc ggacacggct gtaagtcttg tgaatttcat cgcatgaata      1080
ctggagataa gaatgtgatg tgtgcctgtg gctatatgag agcttacaac cattgtgttt      1140
acagtaagtg tgattaactt tagctgggga ggcagagggt gactgggtgc tgactgtttt      1200

```

-continued

---

|   |      |
|---|------|
| atztatgtat atgtttttta tgtgtaggtc ccgctcttga cgtagatgag acccccacta   | 1260 |
| cagagtgcac ttcatacacc ccagaaattg gcgaggaacc gccccaagat attattcata   | 1320 |
| gaccagttgc agtgagagtc accggggcgt gagcagctgt ggagagtttg gatgacttgc   | 1380 |
| tacaggggtg ggatgaacct ttggacttgt gtaccgggaa acgccccagg cactaagtgc   | 1440 |
| cacacatgtg tgtttactta aggtgatgtc agtatattata ggggtgtggag tgcaataaaa | 1500 |
| tccgtgttga ctttaagtgc gtggtttatg actcaggggt ggggactgtg ggtatataag   | 1560 |
| cagggtcaga cctgtgtggt cagttcagag caggactcat ggagatcttg acagtcttgg   | 1620 |
| aagactttca ccagactaga cagctgctag agaactcacc ggaggagtc tcttacctgt    | 1680 |
| ggagattctg cttcgtgtgg cctctagcta agctagtcta tagggccaag caggattata   | 1740 |
| aggatcaatt tgaggatatt ttgagagagt gtcctggtat ttttgactct ctcaacttgg   | 1800 |
| gccatcagtc tcaacttaac cagagtattc tgagagccct tgacttttct actcctggca   | 1860 |
| gaactaccgc cgcggtagcc ttttttgctt ttatccttga caaatggagt caagaaacc    | 1920 |
| atctcagcag ggattaccgt ctggactgct tagcagtagc tttgtggaga acatggaggt   | 1980 |
| gccagcgcct gaatgcaatc tccggctact tgccagtaca gccggtagac acgctgagga   | 2040 |
| tcctgagtct ccagtcaccc caggaacacc aacgccgcca gcagccgacg caggagcagc   | 2100 |
| agcaagagga ggaccgagaa gagaacctga gagccggtct ggaccctccg gtggcggagg   | 2160 |
| aggaggagta gctgacttgt ttcccagact gcgcccgggt ctgactaggt cttccagtgg   | 2220 |
| acgggagagg gggattaagc gggagaggca tgaggagact agccacagaa ctgaactgac   | 2280 |
| tgatcagctg atgagtcgca gccgcccaga atcgggtgtg tggcatgagg tgcagtcgca   | 2340 |
| gggatagat gaggtctcag tgatgcatga gaaatattcc ctagaacaag tcaagacttg    | 2400 |
| ttggttgag cccgaggatg attgggaggt agccatcagg aattatgcca agctggetct    | 2460 |
| gaggccagac aagaagtaca agattaccaa actgattaat atcagaaatt cctgctacat   | 2520 |
| ttcagggat ggggccgagg tggagatcag taccagagag aggtggcctc tcagatgctg    | 2580 |
| catgatgaat atgtaccggg ggggtgtggg catggaggga gtcacctta tgaacgcgag    | 2640 |
| gttcaggggt gatgggtata atgggtgtgt ctttatggcc aacaccaagc tgacagtgca   | 2700 |
| cggatgctcc ttctttgctt tcaataacat gtgcattgag gcctggggca gtgtttcagt   | 2760 |
| gaggggatgc agtttttcag ccaactgcat gggggtcgtg gccagaacca agagcatggt   | 2820 |
| gtcagtgaag aaatgcctgt tcgagaggtg ccacctgggg gtgatgagcg agggcgaagc   | 2880 |
| caaagtcaaa cactgcgcct ctaccgagac gggctgcttt gtactgatca agggcaatgc   | 2940 |
| caaagtcaag cataatatga tctgtggggc ctcggatgag cgcggctacc agatgctgac   | 3000 |
| ctgcgcgggt gggaacagcc atatgctagc caccgtgcat gtggcctcgc acccccgcaa   | 3060 |
| gacatggccc gagttcgagc acaacgtcat gaccgcctgc aatgtgcacc tgggtccccg   | 3120 |
| ccgaggcatg ttcataccct accagtgcaa catgcaatth gtgaagggtc tgctggagcc   | 3180 |
| cgatgccatg tccagagtga gcctgacggg ggtgtttgac atgaatgtgg agctgtggaa   | 3240 |
| aattctgaga tatgatgaat ccaagaccag gtgcccggcc tgcgaatgag gaggcaagca   | 3300 |
| cgccaggcct cagcccgtgt gtgtggaggt gacggaggac ctgcgacccg atcatttgg    | 3360 |
| gttctcctgc aacgggacgg agttcggctc cagcggggaa gaatctgact agagtgagta   | 3420 |
| gtgtttggga ctgggtggga gcctgcatga tgggcagaat gactaaaatc tgtgtttttc   | 3480 |

---

-continued

---

tgcgcagcag catgagcggg agcgcctcct ttgagggagg ggtattcagc ccttatctga 3540  
cggggcgtct cccctcctgg gcgggagtgc gtcagaatgt gatgggatcc acgggtggacg 3600  
gccggcccgt gcagcccgcg aactcttcaa ccctgacctg cgcgacctg agctcctcgt 3660  
ccgtggacgc agctgcccgc gcagctgtg cttccgcgcg cagcgcctg cgcggaatgg 3720  
ccctggggcg cggctactac agctctctg tggccaactc gagttccacc aataatcccg 3780  
ccagcctgaa cgaggagaag ctgctgctgc tgatggccca gctcgaggcc ctgacctcagc 3840  
gcctggggcg gctgacctg caggtggctc agctgcaggc ggagacgcgg gccgcggtt 3900  
ccacggtgaa aaccaataa aaaatgaatc aataaataa cggagacggt tgttgatttt 3960  
aacacagagt cttgaatctt tatttgattt ttcgcgcgcg gtaggccctg gaccaccggt 4020  
ctcgatcatt gagcaccocg tggatctttt ccaggacctg gtagaggtgg gcttgatgt 4080  
tgaggtacat gggcatgagc ccgtcccggg ggtggaggtg gctccattgc agggcctcgt 4140  
gctcgggggt ggtgtttaa atcacccagt catagcaggg gcgcaggcg tgggtgtgca 4200  
cgatgtcctt gaggaggaga ctgatggcca cgggcagccc cttggtgtag gtgttgacga 4260  
acctgttgag ctgggaggga tgcagcggg gggagatgag atgcatctt gcttgatct 4320  
tgagattggc gatgttccc cccagatccc gccgggggtt catgttgtgc aggaccacca 4380  
gcacggtgta tccggtgcac ttggggaatt tgcagcaaa cttggaaggg aaggcgtgaa 4440  
agaatttga gacgccctt tgaccgccca ggttttccat gcaactcatcc atgatgatgg 4500  
cgatgggccc gtgggcggcg gcttgggcaa agacgtttcg ggggtcggac acatcgtagt 4560  
tgtgtcctg ggtgagctc tcataggcca ttttaataa tttggggcg aggggtgccg 4620  
actgggggac gaaggtgcc tcgatcccgg gggcgtagt gccctcgcag atctgcatct 4680  
cccaggcctt gagctcggag ggggggatca tgtccacctg cggggcgatg aaaaaacgg 4740  
tttcggggc gggggagatg agctggggcg aaagcaggt cgggagcagc tgggacttgc 4800  
cgcagccggt ggggcgtag atgaccccga tgaccggctg caggtggtg ttgagggaga 4860  
gacagctgcc gtccctcgcg aggagggggg ccacctcgtt catcatctcg cgcacatgca 4920  
tgttctcgcg cagagttcc gccaggagc gctcgcctcc aagcagagag agctcttgca 4980  
gcgagggcaa gtttttcagc gcttgagcc cgtcggccat gggcatttt gagaggtct 5040  
gttgcaagag ttccagacgg tcccagagct cggatgatgt ctctagggca tctcgatcca 5100  
gcagacctcc tcgtttcgcg ggttggggcg actcggggag tagggcacca ggcgatgggc 5160  
gtccagcag gccaggttcc ggtccttcca ggggcgcagg gtcgcgctca gcgtggtctc 5220  
cgtcacggtg aaggggtgcg cgcgggctg ggcgcttgc aggggtgcgt tcaggctcat 5280  
ccggctggtc gagaaccgct cccggtcggc gccctgcgcg tcggccaggt agcaattgag 5340  
catgagttcg tagttgagc cctcggcccg gtggccctt gcgcggagct tacctttgga 5400  
agtgtgtccg cagacgggac agaggaggga cttgagggcg tagagcttg gggcgaggaa 5460  
gacggactcg gggcgtagg cgtccgcgcg gcagctggcg cagacggtct cgcactccac 5520  
gagccaggtg aggtctggcc ggtcggggtc aaaaacgagg tttcctcctg gctttttgat 5580  
gcgtttctta cctctggtct ccatgagctc gtgtccccg tgggtgacaa agaggctgtc 5640  
cgtgtcccc tagaccgact ttatggggcg gtcctcagc ggggtgcgcg ggtcctcgtc 5700  
gtagaggaac cccgcccact ccgagacgaa ggcccgggtc caggccagca cgaaggaggc 5760

-continued

---

|             |            |             |             |            |             |      |
|-------------|------------|-------------|-------------|------------|-------------|------|
| cacgtgggag  | gggtagcgg  | cggtgtccac  | cagcgggtcc  | accttctcca | gggtatgcaa  | 5820 |
| gcacatgtcc  | ccctcgtcca | catccaggaa  | ggtgattggc  | ttgtaagtgt | aggccacgtg  | 5880 |
| accggggggtc | ccggccgggg | gggtataaaa  | gggggcgggc  | ccctgctcgt | cctcaactgtc | 5940 |
| ttccggatcg  | ctgtccagga | gcgccagctg  | ttggggtagg  | tattccctct | cgaaggcggg  | 6000 |
| catgacctcg  | gcactcaggt | tgctcagttc  | tagaaacgag  | gaggatttga | tattgacggt  | 6060 |
| gccgttgagg  | acgcctttca | tgagcccctc  | gtccatctgg  | tcagaaaaga | cgatcttttt  | 6120 |
| gttgtcgagc  | ttggtggcga | aggagccgta  | gagggcgttg  | gagagcagct | tggcgatgga  | 6180 |
| gcgcgatggtc | tggttctttt | ccttgctggc  | gcgctccttg  | gcggcgatgt | tgagctgcac  | 6240 |
| gtactcgcgc  | gccacgcact | tccattcggg  | gaagacggtg  | gtgagcttgt | cgggcacgat  | 6300 |
| tctgacctcg  | cagcccggtt | tgctcagggg  | gatgaggtcc  | acgctggtgg | ccacctcgcc  | 6360 |
| gcgcaggggc  | tcggttggtc | agcagaggcg  | cccgcccttg  | gcgagcaga  | agggggcgag  | 6420 |
| cgggtccagc  | atgagctcgt | cgggggggtc  | ggcgtccacg  | gtgaagatgc | cgggcaggag  | 6480 |
| ctcggggtcg  | aagtactgta | tgccaggtgc  | cagatcgtcc  | agcgcctctt | gccagtcgcg  | 6540 |
| cacggccagc  | gcgcgctcgt | aggggctgag  | gggctgccc   | cagggcatgg | ggtgcgtgag  | 6600 |
| cgcggaggcg  | tacatgccgc | agatgtcgtg  | gacgtagagg  | ggctcctcga | ggacgccgat  | 6660 |
| gtaggtgggg  | tagcagcgcc | cccccggtat  | gctggcgcgc  | acgtagctgt | acagctcgtg  | 6720 |
| cgagggcgcg  | aggagcccg  | tgccgaggtt  | ggagcgtgc   | ggcttttcgg | cgcggtagac  | 6780 |
| gatctggcgg  | aagatggcgt | gggagttgga  | ggagatggtg  | ggcctctgga | agatgtttaa  | 6840 |
| gtgggcgtgg  | ggcagtcgga | ccgagtcctt  | gatgaagtgg  | gcgtaggagt | cctgcagctt  | 6900 |
| ggcgacgagc  | tcggcgggtg | cgaggacgtc  | cagggcgag   | tagtcgaggg | tctcttggtat | 6960 |
| gatgtcgtac  | ttgagctggc | ccttctgctt  | ccacagctcg  | cggttgagaa | ggaactcttc  | 7020 |
| gcggctcttc  | cagtactctt | cgaggggaa   | cccgtcctga  | tcggcacggt | aagagcccac  | 7080 |
| catgtagaac  | tggttgacgg | ccttgtaggc  | gcagcagccc  | ttctccaacg | ggagggcgta  | 7140 |
| agcttgccgc  | gccttgccga | gggaggtgtg  | ggtgagggcg  | aaggtgtcgc | gcaccatgac  | 7200 |
| cttgaggaac  | tgggtcctta | agtcgaggtc  | gtcgcagccg  | ccctgctccc | agagctggaa  | 7260 |
| gtcctgctgc  | ttcttgtagg | cggggttggg  | caaagcgaag  | gtaacatcgt | tgaagaggat  | 7320 |
| cttgcccgcg  | cggggcatga | agttgcgagt  | gatgcgaaa   | ggctggggca | cctcggcccg  | 7380 |
| gttgttgatg  | acctggggcg | cgaggacgat  | ctcgtcgaag  | ccgttgatgt | tgtgcccgac  | 7440 |
| gatgtagagt  | tccacgaatc | gcgggcggcc  | cttgacgtgg  | ggcagcttct | tgagctcgtc  | 7500 |
| gtaggtgagc  | tcggcggggg | cgctgagggc  | gtgctgctcg  | aggcccaggt | cggcgagggtg | 7560 |
| gggggtggcg  | ccgaggaagg | aagtccagag  | atccacggcc  | aggcgggtct | gcaagcggtc  | 7620 |
| ccggtactga  | cggaaactgt | ggcccacggc  | catttttttcg | ggggtgacgc | agtagaaggt  | 7680 |
| gcgggggtcg  | ccgtgccagc | ggtcccactt  | gagctggagg  | gcgaggtcgt | gggcgagctc  | 7740 |
| gacgagcggc  | gggtccccgg | agagtttcat  | gaccagcatg  | aaggggacga | gctgcttgcc  | 7800 |
| gaaggacccc  | atccaggtgt | aggtttccac  | gtcgtaggtg  | aggaagagcc | tttcggtgcg  | 7860 |
| aggatgcgag  | ccgatgggga | agaactggtat | ctcctgccac  | cagttggagg | aatggctggt  | 7920 |
| gatgtgatgg  | aagtagaagt | gccgacggcg  | gcgccgagc   | tcgtgcttgt | gtttatataa  | 7980 |
| gcgtccgcag  | tgctcgaac  | gctgcacggg  | atgcacgtgc  | tgacagagct | gtacctgggt  | 8040 |



---

-continued

---

tcctttgacg aggaatttca gtgggcagtg gagcgctggc ggctgcatct ggtgctgtac 8100  
tacgtcctgg ccatcgccgt ggccatcgtc tgcctcgatg gtggtcatgc tgacgaggcc 8160  
gcgcgggagg caggccaga cctcggctcg gacgggctcg agagcgagga cgagggcgcg 8220  
caggccggag ctgtccaggg tcctgagacg ctgcgagtc aggtcagtgg gcagcggcgg 8280  
cgcgcggttg acttgacgga gcttttccag ggcgcgcggg aggtccagat ggtacttgat 8340  
ctccacggcg ccggttggtg cgacgtccac ggcttgacag gtcccgtgcc cctggggcgc 8400  
caccaccgty ccccgtttct tcttggtgtc tggcgggcgc ggctccatgc ttagaagcgg 8460  
cggcgaggac gcgcgcggg cggcaggggc ggctcggggc ccgagggcag gggcggcagg 8520  
ggcacgtcgg cgcgcgcgc gggcaggttc tggtagctgc cccggagaag actggcgtga 8580  
gcgacgacgc gacggttgac gtcctggatc tgacgcctct ggtggaaggc cacgggaccc 8640  
gtgagtttga acctgaaaga gagttcgaca gaatcaatct cggtatcgtt gacggcggcc 8700  
tgccgcagga tctcttgac gtcgcccag ttgtcctggt aggcgatctc ggtcatgaac 8760  
tgctcgatct cctcctcctg aaggtctcgg cgaccggcgc gctcgacggt ggcgcgagg 8820  
tcgttgaga tgcggccat gagctgagag aaggcgttca tgcggcctc gttccagacg 8880  
cggctgtaga ccacggctcc gtcgggctcg cgcgcgcgca tgaccacctg ggcgaggttg 8940  
agctcgacgt ggcgctgaa gaccgcgtag ttgcagaggc gctgtagag gtagttgagc 9000  
gtggtggcga tgtgctcgt gacgaagaag tacatgatcc agcggcggag cggcatctcg 9060  
ctgacgtcgc ccaggcttc caagcgtcc atggcctcgt agaagtccac ggcgaagttg 9120  
aaaaactggg agttgcgcgc cgagacggtc aactcctcct ccagaagacg gatgagctcg 9180  
gcgatggtg cgcgcacctc gcgctcgaag gccccggggg gctcctcttc ttccatctcc 9240  
tcctcctctt ccatctcctc cactaacatc tcttctactt cctcctcagg aggcggcggc 9300  
gggggagggg cctgctgctg ccggcggcgc acgggcagac ggtcgatgaa gcgctcgatg 9360  
gtctccccgc gccggcgacg catggtctcg gtgacggcgc gccctcctc gcggggcgcg 9420  
agcgtgaaga cgcgcgcgc catctccagg tggccgcgg gggggtctcc gttgggcagg 9480  
gagagggcgc tgacgatgca tcttatcaat tggcccgtag ggactccgcg caaggacctg 9540  
agcgtctcga gatccacggg atccgaaaac cgctgaacga aggccttcgag ccagtcgcag 9600  
tcgcaaggta ggctgagccc ggtttcttgt tcttcgggta tttggtcggg aggcggcggg 9660  
gcgatgctgc tggtagatgaa gttgaagtag gcggtcctga gacggcggat ggtggcggag 9720  
agcaccaggt ccttgggccc ggcttctggt atgacgagac ggtcggccat gccccaggcg 9780  
tggctctgac acctggcgag gtccttgtag tagtctgca tgagccgctc cacgggcacc 9840  
tcctcctcgc ccgcgcggcc gtgcatgccc gtgagcccga acccgcgctg cggctggacg 9900  
agcgcaggt cggcgacgac gcgctcggcg aggatggcct gctggatctg ggtgagggtg 9960  
gtctggaagt cgtcgaagtc gacgaagcgg tggtaggctc cgggtgtgat ggtgtaggag 10020  
cagttggcca tgacggacca gttgacggtc tggtagcggg ggcgcacgag ctctgtgtac 10080  
ttgagggcgc agtagggcgc cgtgtogaag atgtagtctg tgacggctgc cacgaggtac 10140  
tggtagcga cgaggaagt cggcgggcgc tggcggtaga gcggccatcg ctcggtggcg 10200  
ggggcggcg gcgcgaggtc ctcgagcatg aggcgggtgt agcctgatg gtacctggac 10260  
atccaggtga tgccggcggc ggtggtggag gcgcgggga actcgggac gcggttccag 10320

---

-continued

---

atgttgcgca gcggcaggaa gtagttcatg gtggcccgcg tctggcccgt gaggcgcgcg 10380  
cagtcgtgga tgctctagac atacggggcaa aaacgaaagc ggtcagcggc tcgactccgt 10440  
ggcctggagg ctaagcgaaac gggttgggct gcgcgtgtac cccggttcga gtccttctc 10500  
gaatcaggct ggagccgcag ctaacgtggt actggcactc ccgtctcgac ccaagcctgc 10560  
taacgaaacc tccaggatac ggaggcgggt cgttttgcc atttctgca ggcgggaaat 10620  
gaaactagta agcgcgaaa gcggccgtcc gcgatggctc gctgccgtag tctggagaaa 10680  
gaatcgccag gtttgcttg cgggtgcccc cggttcgagc ctcagcctc ggcgcccggc 10740  
ggattccgcy gctaactgtg gcgtggtgc cccgtcgtt ccaagacccc ttagccagcc 10800  
gacttctcca gttacggagc gagccctct ttttcttggt ttttgccag atgcatccc 10860  
tactgcggca gatgcgcccc caccctccac cacaaccgcc cctaccgcag cagcagcaac 10920  
agccggcgtc tctgccccgc cccagcagc agcagccagc cactaccgcy gcggcccgcg 10980  
tgagcggagc cggcgttcag tatgacctg ccttggaaaga gggcgagggg ctggcgcggc 11040  
tgggggcgtc gtcgcccggag cggcaccgc gcgtgcagat gaaaaggagc gctcgcgagg 11100  
cctacgtgcc caagcagaac ctggtcagag acaggagcgg cgaggagccc gaggagatgc 11160  
gcgcctccc cttccacgcg gggcgggagc tgcggcgcgg cctggaccga aagcgggtgc 11220  
tgagggacga ggatttcgag gcggacgagc tgacggggat cagccccgcg cgcgcgcacg 11280  
tgcccgcgcc caacctggtc acggcgtacg agcagaccgt gaaggaggag agcaacttcc 11340  
aaaaatcctt caacaaccac gtgcgcacgc tgatcgcgcy cgaggaggtg accctgggccc 11400  
tgatgcacct gtgggacctg ctggaggcca tcgtgcagaa ccccacgagc aagccgctga 11460  
cggcgcagct gtttctggtg gtgcagcaca gtcgggacaa cgagacgttc agggaggcgc 11520  
tgctgaatat caccgagccc gagggccgct ggctcctgga cctggtgaac attctgcaga 11580  
gcatcgtggt gcaggagcgc gggctgccgc tgtccgagaa gctggcggcc atcaacttct 11640  
cgggtgctgag cctgggcaag tactacgcta ggaagatcta caagaccccg tacgtgccc 11700  
tagacaagga ggtgaagatc gacgggtttt acatgcgcat gaccctgaaa gtgctgacct 11760  
tgagcagca tctgggggtg taccgcaacg acaggatgca ccgcgcggtg agcgcagcc 11820  
gccggcgcga gctgagcagc caggagctga tgcacagcct gcagcgggccc ctgaccgggg 11880  
ccgggaccga ggggagagc tactttgaca tggcgcggga cctgcgctgg cagcctagcc 11940  
gccgggccc t ggaagctgcc ggcggttccc cctacgtgga ggagggtgac gatgaggagg 12000  
aggaggcgca gtacctgaa gactgatggc gcgaccgtat ttttgctaga tgcagcaaca 12060  
gccaccgccc cctcctgac cgcgatgcy ggcggcgcg cagagccagc cgtccgcat 12120  
taactcctcy gacgattgga cccaggccat gcaacgcatc atggcgtga cgaccgcaa 12180  
tccogaagcc tttagacagc agcctcagc caaccgactc tcggccatcc tggaggccgt 12240  
ggtgccctcy cgtcgaacc ccacgcagca gaaggtgctg gccatcgtga acgcgctggt 12300  
ggagaacaag gccatccgcy gcgacgagc cgggctggtg tacaacgcgc tgcgggagcy 12360  
cgtggcccgc tacaacagca ccaacgtgca gacgaacctg gaccgatgg tgaccgagct 12420  
gcgcgagggc gtgtcgcagc gcgagcgggt ccaccgcgag tcgaaacctg gctccatggt 12480  
ggcgtgaac gccttctga gcacgcagcc cgcacaactg ccccggggccc aggaggacta 12540  
caccaacttc atcagcgcgc tgcggctgat ggtggccgag gtgcccaga gcgaggtgta 12600

---

-continued

---

ccagtcgggg ccggactact tcttccagac cagtcgccag ggcttgcaga ccgtgaacct 12660  
gagccaggct ttcaagaact tgcagggact gtggggcgtg caggccccgg tcggggaccg 12720  
cgcgacggtg tcgagcctgc tgacgccaa ctcgcgcctg ctgctgctgc tggtagcgcc 12780  
cttcacggac agcggcagcg tgagccgca ctcgtacctg ggctacctgc ttaacctgta 12840  
cccgagggcc atcgggcagc cgcacgtgga cgcgcagacc taccaggaga tcaaccacgt 12900  
gagccgcgcg ctggggcagg aggacccggg caacctggag gccaccctga acttcctgct 12960  
gaccaaccgg tcgcagaaga tcccgccca gtacgcgctg agcaccgagg aggagcgcat 13020  
ctcgcgctac gtgcagcaga gcgtggggct gttcctgatg caggaggggg ccacgcccag 13080  
cgccgcgctc gacatgaccg cgcgcaacat ggagcccagc atgtacgcc gcaaccgccc 13140  
gttcacatcaat aagctgatgg actacttgca tcggggcgcc gccatgaact cggactactt 13200  
taccaacgcc atcttgaacc cgcactggct cccgccgccc gggttctaca cgggcgagta 13260  
cgacatgccc gaccccaacg acgggttctc gtgggacgac gtggacagca gcgtgttctc 13320  
gcccgcgccc accaccacca ccgtgtgaa gaaagaggc ggggaccggc ggcgctctc 13380  
ggcgtgtcc ggtcgcgcg gtgctgccg ggcggtgccc gagccgcca gcccttccc 13440  
gagcctgccc tttctgctga acagcgtgcg cagcagcgag ctgggtcggc tgacgcgcc 13500  
gcccctgctg ggcgaggagg agtacctgaa cgactccttg cttcggcccg agcgcgagaa 13560  
gaacttcccc aataacggga tagagagcct ggtggacaag atgagccgct ggaagacgta 13620  
cgcgcacgag cacagggacg agccccgagc tagcagcagc accggcgcga cccgtagacg 13680  
ccagcggcac gacaggcagc ggggtctggt gtgggacgat gaggattccg ccgacgacag 13740  
cagcgtgttg gacttggtg ggagtgtggt tggtaacccg ttcgctcacc tgcgccccg 13800  
tatcgggccc ctgatgtaag aatctgaaaa aataaaagac ggtactcacc aaggccatgg 13860  
cgaccagcgt gcgttcttct ctggtgtttg tagtagtatg atgaggcgcg tgtaccggga 13920  
gggtcctcct ccctcgtacg agagcgtgat gcagcagcg gtggcggcgg cgatgcagcc 13980  
cccgtggag gcgccttacg tgcccccggt gtacctggcg cctacggagg ggcggaacag 14040  
cattcgttac tcggagctgg cacccttcta cgataccacc cggttgtacc tggtaggaaa 14100  
caagtcggcg gacatcgcct cgctgaacta ccagaacgac cacagcaact tcctgaccac 14160  
cgtggtgcag aacaacgatt tcacccccac ggaggccagc acccagacca tcaactttga 14220  
cgagcgcctc cggtagggcg gccagctgaa aaccatcatg cacaccaaca tgcccaacgt 14280  
gaacgagttc atgtacagca acaagttcaa ggcgcggtg atggtctcgc gcaagacccc 14340  
caacggggtc acagtaacag atggtagtca ggacgagctg acctacgagt gggtaggatt 14400  
tgagctgccc gagggcaact tctcgggtgac catgaccatc gatctgatga acaacgccat 14460  
catcgacaac tacttgccg tggggcggca gaacggggtg ctggagagcg acatcggcgt 14520  
gaagttcgac acgcgcaact tccggtggg ctgggacccc gtgaccgagc tggtagtgcc 14580  
ggcgtgtgac accaacgagg ccttccaccc cgacatcgtc ctgctgcccg gctgcggcgt 14640  
ggacttcacc gagagccgcc tcagcaacct gctgggcatc cgcaagcggc agccttcca 14700  
ggagggcttc cagatcctgt acgaggacct ggaggggggc aacatccccg cgctgctgga 14760  
cgtgagcgc tacgagaaaa gcaaggagga tagcgcgccc gcggcgaccg cagccgtggc 14820  
caccgcctct accgaggtgc ggggcgataa ttttgctagc gccgcgacac tggcagcggc 14880

---

-continued

---

cgaggcggct gaaaccgaaa gtaagatagt gatccagccg gtggagaagg acagcaagga 14940  
gaggagctac aacgtgctcg cggacaagaa aaacaccgcc taccgcagct ggtacctggc 15000  
ctacaactac ggcgaccccc agaagggcgt gcgctcctgg acgctgctca ccacctcgga 15060  
cgtcacctgc ggcgtggagc aagtctactg gtcgctgccc gacatgatgc aagaccgggt 15120  
caccttccgc tccacgcgtc aagttagcaa ctacccggtg gtggggcggc agctcctgcc 15180  
cgtctactcc aagagcttct tcaacgagca ggccgtctac tcgcagcagc tgcgcgcctt 15240  
cacctcgctc acgcacgtct tcaaccgctt ccccgagaac cagatcctcg ttcgcccggc 15300  
cgcgcccacc attaccaccg tcagtgaaaa cgttcctgct ctcacagatc acgggaccct 15360  
gccgctgctc agcagtatcc ggggagtcca gcgcgtgacc gtcactgacg ccagacgccc 15420  
cacctgcccc tacgtctaca aggcctggg cgtagtgcgc ccgcgcgtcc tctcgagccg 15480  
caccttctaa aaaatgtcca ttctcatctc gcccagtaat aacaccgggt ggggcctgct 15540  
cgcgcccagc aagatgtacg gaggcgctcg ccaacgctcc acgcaacacc ccgtgcgcgt 15600  
gcgcgggac ttccgcgctc cctggggcgc cctcaagggc cgcgtgcgct cgcgcaccac 15660  
cgtgcagcag gtgatcgacc aggtggtggc cgacgcgcgc aactacacgc ccgcccgccg 15720  
gcccgtctcc accgtggagc ccgtcatcga cagcgtggtg gccgacgcgc gccggtacgc 15780  
ccgcgccaaag agccggcggc gccgcacgc ccggcgggc cggagcacc ccgccatgct 15840  
cgcggcgca gccttctgct gcaggggcag gcgcacggga cgcaggggca tgcctcagggc 15900  
ggccagacgc gcggcctccg gcagcagcag gcggcgagc acccgacagc gcgcggccac 15960  
ggcgcgggcg gcggccatcg ccagcatgct ccggcccgcg cgcggcaacg tgtactgggt 16020  
gcgcgacgcc gccaccgggt tgcgcgtgcc cgtgcgcacc cgcggccctc gcacttgaag 16080  
atgctgactt cgcgatggtg atgtgtccca gcgcgcgagga ggtgtccaa gcgcaaattc 16140  
aaggaaagaga tgctccaggt catcgcgcct gagatctacg gcccgggcggc ggtgaaggag 16200  
gaaagaaagc cccgcaaact gaagcgggtc aaaaaggaca aaaaggagga ggaagatgtg 16260  
gacggactgg tgaggtttgt gcgcgagttc gcccccggc ggcgctgca gtggcgcggg 16320  
cggaaagtga aaccggtgct gcgacccgac accacgggtg tcttcacgcc cggcgagcgt 16380  
tccggtccg cctccaagcg ctccctacgac gaggtgtacg gggacgagga catcctcgag 16440  
caggcgggcc aacgtctggg cgagtttctg tacggcaagc gcagccgccc cgcggccttg 16500  
aaagaggag cgggtgccat cccgctggac cacggcaacc ccacgcccag cctgaagccg 16560  
gtgaccctgc agcaggtgct gcctggtgct gcgcgcgcc ggggcttcaa gcgcgagggc 16620  
ggcgaggatc tgtaccgcac catgcagctg atggtgcccc agcgccagaa gctggaggac 16680  
gtgctggagc acatgaaggt ggaccccgag gtgcagccc aggtcaaggt gcggcccatc 16740  
aagcaggtgg ccccgggcct gggcgtgac accgtggaca tcaagatccc cacggagccc 16800  
atggaacgc agaccgagcc cgtgaagccc agcaccagca ccatggaggt gcagacggat 16860  
ccctggtgac cggcaccggc ttccaccacc cgcggaagac gcaagtacgg cgcggccagc 16920  
ctgtgatgc ccaactacgc gctgcatcct tccatcatcc ccacgcccgg ctaccgcccg 16980  
acgcgcttct accgcggcta caccagcagc gcggccgca agaccaccac ccggcccgcc 17040  
cgtcgtcgca cccgcccagc cagcaccgag acttccgccc ccggcctggg gcgggaggtg 17100  
taccgcagc gcgcgagcc tctgaacctg ccgcgcgcgc gctaccaacc gagcatcgcc 17160

---

-continued

---

atttaactac cgctcctac ttgcagatat ggccctcaca tgccgcctcc gcgtcccat 17220  
tacgggctac cgaggaagaa agccgcgccc tagaaggctg acggggaacg ggctgcgctg 17280  
ccatcaccac cggcggcggc gcgccatcag caagcggttg gggggaggct tcctgcccgc 17340  
gctgatgccc atcatcgccc cggcgatcgg ggcgatcccc ggcatagctt ccgtggcggg 17400  
gcaggcctct cagcgccact gagacacagc ttggaaaatt tgtaataaaa aatggactga 17460  
cgctcctggt cctgtgatgt gtgtttttag atggaagaca tcaatttttc gtcctggca 17520  
ccgcgacacg gcacgcggcc gtttatgggc acctggagcg acatcgcaa cagccaactg 17580  
aacggggggc cttcaattg gagcagctc tggagcggg ttaagaattt cgggtccacg 17640  
ctcaaacct atggcaacaa ggcgtggaac agcagcacag ggcaggcgcct gagggaaaag 17700  
ctgaaagagc agaacttcca gcagaagggt gtcgatggcc tggcctcggg catcaacggg 17760  
gtggtggacc tggccaacca ggcctgcag aacagatca acagccgctt ggacgcggtc 17820  
ccgcccgcgg ggtccgtgga gatgcccag gtggaggagg agctgcctcc cctggacaag 17880  
cgcggcgaca agcagaccgc tcccagcgc gaggagacgc tgctgacgca cagggacgag 17940  
ccgccccgt acgaggaggc ggtgaaactg ggtctgccc ccacgcggcc cgtggcgcct 18000  
ctggccaccg ggtgctgaa acccagcagc agcagcagcc agcccgcgac cctggacttg 18060  
ctccgcctg cttcccggcc ctccacagt gctaagcccc tgccgcgggt ggcctgcg 18120  
tcgcgcgccc cccgaggccc ccccaggcg aactggcaga gactctgaa cagcatcgtg 18180  
ggtctgggag tgcagagtgt gaagcggcc cgctgctatt aaaagacact gtagcgctta 18240  
acttgcttgt ctgtgtgat atgtatgtcc gccgaccaga aggaggagga agaggcgcgt 18300  
cgccgagttg caagatggcc acccaccga tgctgcccc gtggcgctac atgcacatcg 18360  
ccggacagga cgettcggag tacctgagtc cgggtctggt gcagttcgc ccgcccacag 18420  
acacctactt cagtctgggg aacaagtta ggaacccac ggtggcggc acgcacgatg 18480  
tgaccaccga ccgacggcag cggctgacgc tgcgcttctg gccctggac cgcgaggaca 18540  
acacctactc gtacaaagt cgctacacgc tggccgtggg cgacaaccgc gtgctggaca 18600  
tggccagcac ctactttgac atccgcggcg tgctggatcg gggccctagc ttcaaacct 18660  
actccggcac cgttacaac agcctggctc ccaaggagc gcccaacact tgccagtgga 18720  
catataaagc tgatggtgat actggtacag aaaaaaccta tacatatgga aatgcgctg 18780  
tgcaaggcat tagtattaca aaagatgta ttcaacttg aactgacact gatgatcagc 18840  
ccatttatgc agataaaact tatcaaccag agcctcaagt ggggtgatgt gaatggcatg 18900  
acatcactgg tactgatgaa aaataggag gcagagctct caagcctgac accaaaatga 18960  
agcctgcta tggttctttt gccaaagccta ccaataaaga aggaggcag gcaaatgtga 19020  
aaaccgaaac aggcgttacc aaagaatag acattgacat ggcattcttc gataatcgaa 19080  
gtgcagctgc ggctggcctg gcccagaaa ttgttttgta tactgagaat gtggatctgg 19140  
aaactcaga tactcatatt gtatacaagg cgggcacaga tgacagcagc tcttctatca 19200  
atgtgggtca gcagtcctag cccaacagac ccaactacat tggctttaga gacaacttta 19260  
tcgggctcat gtactacaac agcactggca acatggcgt gctggctggg caggcctccc 19320  
agctgaatgc tgtggtggac ttgcaggaca gaaacactga actgtcctac cagctcttgc 19380  
ttgactctct gggcgacaga accaggtatt tcagtatgtg gaatcaggcg gtggacagct 19440

---

-continued

---

atgaccccca tgtgcccatt attgaaaac acggtgtgga ggatgaactc cctaactatt 19500  
gcttccccct ggatgctgtg ggtagaactg ataactacca gggaaattaag gccaatggtg 19560  
ctgatcaaac cacctggacc aaagatgata ctgttaatga tgctaataaa ttgggcaagg 19620  
gcaatccttt cgccatggag atcaacatcc aggccaacct gtggcggaac ttctctacg 19680  
cgaacgtggc gctgtacctg cccgactcct acaagtaac gccggccaac atcacgctgc 19740  
cgaccaaac caacacctac gattacatga acggccgctg ggtggcgccc tcgctggtgg 19800  
acgctcatat caacatcggg gcgcgctggt cgctggaccc catggacaac gtcaacccct 19860  
tcaaccacca ccgcaacgcg ggcctgcgct accgctccat gctcctgggc aacggcgct 19920  
acgtgccttt ccacatccag gtgccccaaa agttcttcgc catcaagagc ctctgctcc 19980  
tgcccggtc ctacacctac gactggaact tccgcaagga cgcaacatg atcctgcaga 20040  
gctccctcgg caacgacctg cgcacggacg gggcctccat cgccttcacc agcatcaacc 20100  
tctacgccac ctctctcccc atggcgcaaca acaccgctc cagctcgag gccatgctgc 20160  
gcaacgacac caacgaccag tcttcaacg actacctctc ggcggccaac atgctctacc 20220  
ccatcccggc caacgcccac aacgtgccca tctccatccc ctgcgcaac tgggcccct 20280  
tccgcgatg gtccttcacg cgcctcaaga cccgcgagac gccctcgtc ggtcccgggt 20340  
tcgacccta ctctctctac tcgggctcca tcccctacct cgacggcacc ttctacctca 20400  
accacacctt caagaaggtc tccatcacct tcgactctc cgtcagctgg cccggcaacg 20460  
accgctcct gacgccaac gacttcgaaa tcaagcgac cgtcgacgga gaggggtaca 20520  
acgtggccca gtgcaacatg accaaggact ggttctggt ccagatgctg gccactaca 20580  
acatcggtc ccagggttc tacgtgccg agggctacaa ggaccgatg tactcctct 20640  
tccgcaactt ccagcccatg agccgccagg tcgtggacga ggtcaactac aaggactacc 20700  
aggccgtc cctggctac cagcacaaca actcgggctt cgtcggctac ctgcgcca 20760  
ccatgcgcca gggacagccc taccocgcca actacccta cccgtctac ggaagagcg 20820  
ccgtcggcag cgtcaccag aaaaagtcc tctgcgaccg ggtcatgtg cgcacctct 20880  
tctcagcaa ctctatgtc atggcgcgcc tcaccgacct cggccagaac atgctctacg 20940  
ccaactccgc ccacgcgcta gacatgaatt tcgaagtcga cccatggat gactccacc 21000  
ttctctatgt tgtcttoga gtcttcgacg tcgtccgagt gcaccagccc caccgcgcg 21060  
tcacgaggg cgtctacctg cgcacgcct tctcggcgg caacgccacc acctaaagccc 21120  
cgctcttct tcttgcaaga tgacggcctg tgcgggctcc ggcgagcagg agctcagggc 21180  
catcctccgc gacctgggt gcgggcccgt cttcctgggc accttcgaca agcgttccc 21240  
gggattcatg gccccgaca agctggcctg cgcctcgtc aacacggccg gccgagagc 21300  
cggggcgag cactggtg ccttcgctg gaaccgcgc tcccacaact gctacctct 21360  
cgacccctc gggttctcg acgagcgct caagcagatc taccagttcg agtacgagg 21420  
cctgctgcgc cgcagcgcct tggccaccga ggaccgctgc gtcaccctg aaaagtccac 21480  
ccagaccgtg cagggtccgc gctcggcgc ctgcgggctc ttctgctgca tgttctgca 21540  
cgctctcgt cactggccc accgcccct ggacaagaac cccacctga acttgctgac 21600  
gggggtgcc aacggcatg tcagctgcc ccaggtgga cccacctgc gccgcaacca 21660  
ggagcgctc taccgcttc tcaacgcca ctccgctac tttcgtccc accgcgcg 21720

---

-continued

---

catcgagaag gccaccgcct tcgaccgcat gaatcaagac atgtaaaccg tgtgtgtatg 21780  
tgaatgcttt attcataata aacagcacat gtttatgccca ccttttctga ggctctgact 21840  
ttatntagaa atcgaagggg ttctgccggc tctcggcgtg ccccgcgggc agggatacgt 21900  
tgcggaactg gtacttgggc agccacttga actcggggat cagcagcttc ggcacgggga 21960  
ggcgggggaa cgagtgcctc cacagcttgc gcgtgagttg cagggcgccc agcaggtcgg 22020  
gcgcggagat cttgaaatcg cagttgggac ccgcgttctg cgcgcgggag ttgcggtaca 22080  
cggggttga cacttgggac accatcaggg ccgggtgctt cacgctcgcc agcaccgtcg 22140  
cgtcgggtgat gccctccacg tccagatcct cggcgttggc catcccgaag ggggtcatct 22200  
tgcaggtctg ccgccccatg ctgggcacgc agccgggctt gtggttcaa tcgcagtgca 22260  
gggggatcag catcatctgg gcctgctcgg agctcatgcc cgggtacatg gccttcatga 22320  
aagcctccag ctggcgggaa gcctgctcgg ccttgccgcc ctcggtgaa aagaccccgc 22380  
aggacttgct agagaactgg ttggtggcgc agccggcgtc gtgcacgcag cagcgcgcgt 22440  
cgttgttggc cagctgcacc acgctgcgcc cccagcgggt ctgggtgatc ttggcccgggt 22500  
cggggttctc cttcagcgcg cgtgcccgt tctcgtcgc cacatccatc tcgatcgtgt 22560  
gctccttctg gatcatcacg gtcccgtgca ggcatcgcag cttgccctcg gcctcgggtc 22620  
accctgtcag ccacagcgcg cagccgggtc actcccagtt cttgtgggcg atctgggagt 22680  
gcgagtgca gaagccctgc aggaagcggc ccacatcgt ggtcagggtc ttgttctctg 22740  
tgaagtcag cgggatgccg cgggtctcct cgttcacata caggtggcag atgcggcgggt 22800  
acacctgcc ctgctcgggc atcagctgga aggcggactt caggtcgtc tccacgcgggt 22860  
accggtccat cagcagcgtc atgacttcca tggccttctc ccaggccgag acgatcggca 22920  
ggctcagggg gttcttcacc gccgttgc tcttagtcgc cgcgctgag gtcagggggt 22980  
cgttctcgtc cagggctca aacactcgt tgcctcctt ctcggtgatg cgcacggggg 23040  
gaaagctgaa gccacggcc gccagctcct cctcggcctg cctttcgtcc tcgctgtcct 23100  
ggctgatgtc ttgcaaagc acatgcttgg tcttgccggg tttctttttg ggcggcagag 23160  
gcggcgcg agactgctg gcgagcgc agttctcgt caccacgact atttcttctt 23220  
cttggccgtc gtccgagacc acgcggcggg aggcattgct cttctggggc agagggcggg 23280  
gcgacgggct ctcgcgggtc ggcgggcgc tggcagagcc ccttccgct tcgggggtgc 23340  
gtcctggcg gcgctgctt gactgacttc ctccgcggcc ggccattgtg ttctcctagg 23400  
gagcaacaag catggagact cagccatcgt cgcacaacac gccatctgcc cccgccgccc 23460  
ccgacgagaa ccagcagcag aatgaaagct taaccgcccc gccgcccagc cccacctccg 23520  
acgcccccg ggcaccagac atgcaagaga tggaggaatc catcgagatt gacctgggct 23580  
acgtgacgcc cgcggagcag gaggaggagc tggcagcgc cttttcagcc ccggaagaga 23640  
accaccaaga gcagccagag caggaagcag agagcagca gcagcaggct gggctcagc 23700  
atggcgacta cctgagcggg gcagaggagc tgctcatcaa gcacttgcc cgcacaatgca 23760  
tcatcgtcaa ggacgcgctg ctcgaaccgc cagaggtgcc cctcagcgtg gcggagctca 23820  
gccgcgccta cgagcgaac ctcttctcgc cgcgcgtgcc cccaagcgc cagcccaacg 23880  
gcacctgca gcccaaccg cgcctcaact tctaccgggt cttcgcgggt cccgagggcc 23940  
tggccacta ccacctctt ttcaagaacc aaaggatccc cgtctcctgc cgcgcaacc 24000

---

-continued

---

gcacccgcgc cgacgccctg ctcaacctgg gtcccggcgc ccgcctacct gatatacgct 24060  
ccttgaaga ggttcccaag atcttcgagg gtctgggcag cgacgagact cgggccgcga 24120  
acgctctgca aggaagcggg gaggagcatg agcaccacag cgccttggtg gagttggaag 24180  
gcgacaacgc gcgcctggcg gtgctcaagc gcacggtcga gctgaccacac ttgcctacc 24240  
cggcgctcaa cctgcccccc aaggctcatg gcgcccgtcat ggaccagggtg ctcatcaagc 24300  
gcgcctcgcc cctctcggat gaggacatgc aggaccccgga gagctcggac gagggcaagc 24360  
ccgtggtcag cgacgagcag ctggcgcgct ggctgggagc gtagtagcacc ccccagagct 24420  
tggaagagcg gcgcaagctc atgatggccg tggctcctggt gaccgtggag ctggagtgtc 24480  
tgcgcccgtt ctctcgcgac gcagagacc tgcgcaaggc cgaggagaac ctgcaactacc 24540  
tcttcaggca cgggtttgtg cgcacaggct gcaagatctc caacgtggag ctgaccaacc 24600  
tggtctccta catgggcatc ctgcaagaga accgcctggg gcagaacgtg ctgcaacca 24660  
ccctgcgcgg ggaggcccgc cgcgactaca tccgcgactg cgtctacctg tacctctgcc 24720  
acacctggca gacgggcatg ggcgtgtgac agcagtgcct ggaggagcag aacctgaaag 24780  
agctctgcaa gctcctgca gaaacctga aggcctgtg gaccgggttc gacgagcga 24840  
ccaccgcctc ggacctggcc gacctcatct tccccgagcg cctgcgctg acgctgcgca 24900  
acggactgcc cgactttatg agtcaaaaga tgttgcaaaa ctttcgctct ttcactctcg 24960  
aacgctccgg gatcctgccc gccacctgct ccgctgctgc ctcggacttc gtgcccgtga 25020  
ccttccgcga gtgccccccg ccgctctgga gccactgcta cctgctgcgc ctggccaact 25080  
acctggccta cactcggac gtgatcgagg acgtcagcgg cgagggtctg ctcgagtgcc 25140  
actgcccgtg caacctctgc acgcccacac gctccctggc ctgcaacccc cagctgctga 25200  
gcgagaccca gatcatcggc accttcgagt tgcaaggccc cggcgagggc aaggggggtc 25260  
tgaaactcac cccggggctg tggacctcgg cctacttgcg caagttctgtg cccgaggact 25320  
accatccctt cgagatcagg ttctaogagg accaatccca gccgcccaag gccgaactgt 25380  
cggcctgcgt catcaccag ggggccatcc tggccaatt gcaagccatc cagaaatccc 25440  
gccaagaatt tctgctgaaa aagggccacg gggctacct ggacccccag accggagagg 25500  
agctcaaccc cagcttcccc caggatgccc cgaggaagca gcaagaagct gaaagtggag 25560  
ctgcccgcgc cgaggattht ggaggaagac tgggagagca gtcaggcaga ggaggaggag 25620  
atggaagact gggacagcac tcaggcagag gaggacagcc tgcaagacag tctggaagac 25680  
gaggtggagg aggaggcaga ggaagaagca gccgccgcca gaccgtcgtc ctgcccggag 25740  
aaagcaagca gcacggatac catctccgct ccgggtcggg gtcgcgcgca ccgggccac 25800  
agtaggtggg acgagaccgg gcgcttcccg aacccccca cccagaccgg taagaaggag 25860  
cggcagggat acaagctcct gcgggggac aaaaacgcca tcgtctcctg cttgcaagcc 25920  
tgcgggggca acatctcctt cccccccgc tacctgctct tccaccgcg ggtgaacttc 25980  
ccccgaaca tcttgatta ctaccgtcac ctccacagcc cctactactg tttccaagaa 26040  
gaggcagaaa cccagcagca gcagaaaacc agcggcagca gcagctagaa aatccacagc 26100  
ggcggcaggt ggactgagga tcgcagcga cgagccggcg cagaccggg agctgaggaa 26160  
ccggtcttt cccacctct atgcatctt ccagcagagt cgggggcag agcaggaaact 26220  
gaaagtcaag aaccgttctc tgcgctcgt caccgcagc tgtctgtatc acaagagcga 26280



---

-continued

---

agaccaactt cagcgcactc tcgaggacgc cgaggctctc ttcaacaagt actgcgcgct 26340  
cactcttaaa gagtagcccg cgcgcccca cacacggaaa aaggcgggaa ttacgtcacc 26400  
acctgcgccc ttgccccgac catcatcatg agcaaagaga tccccacgcc ttacatgtgg 26460  
agctaccagc cccagatggg cctggccgcc ggcgccgcc aggactactc cccccgatg 26520  
aactggctca gcgccgggcc cgcgatgatc tcacgggtga atgacatccg cgcgcccca 26580  
aaccagatac tcctagaaca gtcagcgatc accgccacgc cccgccatca ccttaatccg 26640  
cgtaattggc ccgcccctt ggtgtaccag gaaattcccc agcccacgac cgtactactt 26700  
ccgcgagacg cccagggcga agtccagctg actaactcag gtgtccagct ggcggcgggc 26760  
gccgcccctg gtcgtcaccg ccccgtcag ggtataaagc ggctggatgat ccgaggcaga 26820  
ggcacacagc tcaacgacga ggtggtgagc tcttcgctgg gtctgcgacc tgacggagtc 26880  
ttccaactcg ccgcatcggg gagatcttcc ttcacgcctc gtcaggccgt cctgactttg 26940  
gagagttcgt cctcgcagcc ccgctcgggt ggcacgcgca ctctccagtt cgtggaggag 27000  
ttactccct cggctactt caacccttc tccggctccc cggccacta cccggacgag 27060  
ttcatccga acttcgacgc catcagcgag tcgggtggacg gctacgattg aatgtcccat 27120  
ggtggcgag ctgacctagc tcggcttcga cacctggacc actgccgcc cttccgctgc 27180  
ttcgtcggg atctcgcga gtttgcttac tttgagctgc ccgaggagca ccctcagggc 27240  
ccggcccacg gagtgcggat catcgtcga gggggcctc actcccact gcttcggatc 27300  
ttcagccagc gaccgatcct ggtcagcgc gagcaaggac agaccctct gaccctgtac 27360  
tgcatctgca accaccccg cctgcatgaa agtctttgtt gtctgctgtg tactgagtat 27420  
aataaaaagt gagatcagc actactccg actcgattgt ggtgttcctg ctatcaaccg 27480  
gtccctgttc ttcaccggga acgagaccga gctccagctt cagtgtgtaagc cccacaagaa 27540  
gtacctcacc tggctgttc agggctcccc gatcgcggtt gtcaaccact gcgacaacga 27600  
cggagtctct ctgagcggcc ccgccaacct tactttttcc acccgagaa gcaagctcca 27660  
gctcttccaa cccttctcc ccgggacctc tcagtgcgtc tcgggacct gccatcacac 27720  
cttccacctg atcccgaata ccacagcgc gctccccgt actaacaacc aaactacca 27780  
ccatcgccac cgtcgcgacc tttctgaatc taactactacc acccacaacc gaggtgagct 27840  
ccgaggtcga ccaacctctg ggatttacta cggcccctgg gaggtggtgg ggttaatagc 27900  
gctaggccta gttgtgggtg gctttttgct tctctgtac ctatacctcc cttgctgttc 27960  
gtaactagtg gtgctgtgtt gctggtttaa gaaatgggga agatcacct agtgagctgc 28020  
ggtgcgctgg tggcgggtgt ggtgttttcg attgtgggac tggcgcgcg ggctgtagtg 28080  
aaggagaag ccgatccctg cttgcatttc aatcccgaca attgccagct ggttttcag 28140  
cccgatggca atcggtcgca ggtgctgatc aagtgcggat gggaatgcga gaaagtgaga 28200  
atcaggtaca ataacaagac tcggaacaat actctcgcgt ccgtgtggca gcccggggac 28260  
cccagtggt acaccgtctc tgtccccgt gctgacggct ccccgcgac cgtgaacaat 28320  
actttcattt ttgcgcacat gtcgacacg gtcattgtgga tgagcaagca gtaogatatg 28380  
tggccccca cgaaggagaa catcgtggtc ttctccatcg cttacagcgc gtcacggcg 28440  
ctaataccg ctatcgtgtg cctgagcatt cacatgctca tcgctattcg ccccagaaat 28500  
aatgccgaaa aagagaaca gccataaac gttttttcac acacctttt cagaccatgg 28560

---

-continued

---

cctctgttaa atttttgctt ttatttgcca gtctcattac tgttataagt aatgagaaac 28620  
tcactattta cattggcact aaccacactt tagacggaat tccaaaatcc tcatgggtatt 28680  
gctattttga tcaagatcca gacttaacta tagaactgtg tggtaacaag gaaaaaata 28740  
caagcattca ttaattaac ttaattgag gagacaattt gaaattaatt aatatcacta 28800  
aagagtatgg aggtatgtat tactatgttg cagaaaataa caacatgcag ttttatgaag 28860  
ttactgtaac taatcccacc acacctagaa caacaacaac caccaccaca aaaactacac 28920  
ctgttaccac tatgcagctc actaccaata acatttttgc catgcgctca atgggtcaaca 28980  
atagcactca acccacccca cccagtgagg aaattcccaa atccatgatt ggcattattg 29040  
ttgctgtagt ggtgtgcatg ttgatcatcg ccttgtgcat ggtgtactat gccttctgct 29100  
acagaaagca cagactgaac gacaagctgg aacacttact aagtgttgaa ttttaatttt 29160  
ttagaacat gaagatccta ggccttttaa tttttctat cattacctct gctctatgca 29220  
attctgacaa tgaggacgtt actgtcgttg tcggaaccaa ttatacactg aaagggtccag 29280  
cgaagggtat gcttctgttg tattgctggt ttggaactga cgagcaacag acagagctct 29340  
gcaatgctca aaaaggcaaa acctcaaatt ctaaaatctc taattatcaa tgcaatggca 29400  
ctgacttagt actgctcaat gtcacgaaag catatgctgg cagctacacc tgccctggag 29460  
atgatactga gaacatgatt ttttacaag tggaaagtgt tgatcccact actcccctc 29520  
caccaccacc aactactcac accacacaca cagaacaac cacagcagag gaggcagcaa 29580  
agttagcctt gcaggccaa gacagtccat ttgttgcat taccctaca cctgatcagc 29640  
ggtgtccggg gctgctgctc agcggcattg tcggtgtgct ttcgggatta gcagtcataa 29700  
tcatctgcat gttcattttt gcttctgctc atagaaggct ttaccgacaa aatcagacc 29760  
cactgctgaa cctctatggt taatttttc cagagccatg aaggcagtta gcaacttagt 29820  
ttttgttct ttgattggca ctgtttttag tgttagcttt ttgaaacaaa tcaatgttac 29880  
tgagggggaa aatgtgacac tggtaggcgt agaggggtgct caaataacca cctggacaaa 29940  
attccatcta gatgggtgga aagaaattg cacctggaat gtcagtactt atacatgtga 30000  
aggagttaat cttaccattg tcaatgtcag ccaaattcaa aagggttggg ttaaagggca 30060  
atctgttagt gttagcaata gtgggtacta taccagcat actcttatct atgacattat 30120  
agttatacca ctgcctacac ctagcccacc tagcactacc acacagacaa cccacactac 30180  
acaaacaacc acatacagta catcaaatca gcctaccacc actacaacag cagagggtgc 30240  
cagctcgtct ggggtccgag tggcattttt gatgttgcc ccatctagca gtcccactgc 30300  
tagtaccat gagcagacta ctgaattttt gtccactgct gagagccaca ccacagctac 30360  
ctcgagtgc ttctctagca ccgccaatct atcctcgtt tcctctacac caatcagtc 30420  
cgctactact cctacccccg ctattctccc cactcccctg aagcaaacag acggcgacat 30480  
gcaatggcag atcacctgc tcattgtgat cgggttggc atcctggccg tgttgcctca 30540  
ctacatcttc tgccgccga ttcccaacgc gcaccgcaag ccggcctaca agcccatcgt 30600  
tgtoggcgag ccggagccgc ttcaggtgga aggggtccta aggaatcttc tcttctctt 30660  
tacagtatgg tgattgaatt atgattccta gacaaatctt gatcactatt cttatctgcc 30720  
tcctccaagt ctgtgccacc ctgctctggt tggccaacgc cagtccagac tgtattggc 30780  
ccttgcctc ctacgtgctc tttgccttca tcacctgcat ctgctgctgt agcatagtct 30840

---

-continued

---

gcctgcttat caccttcttc cagttcattg actggatctt tgtgcgcata gcctacctgc 30900  
gccaccaccc ccagtaccgc gaccagcgag tggcggggct gctcaggatc ctctgataag 30960  
catgcgggct ctgctacttc tcgcgcttct gctgttagtg ctccccgcgc ccgtcgaccc 31020  
ccggaccccc acccagtcctc ccgaggaggt ccgcaaatgc aaattccaag aacctggaa 31080  
attctcaaaa tgctaccgcc aaaaatcaga catgcatccc agctggatca tgatcattgg 31140  
gatcgtgaac attctggcct gcaccctcat ctctttgtg atttaccctt gctttgactt 31200  
tggttggaac tcgccagagg cgctctatct ccgcctgaa cctgacacac caccacagca 31260  
acctcaggca cacgcactac caccaccacc acagcctagg ccacaatata tgcccatatt 31320  
agactatgag gccgagccac agcgaccat gctccccgct attagttact tcaatctaac 31380  
cggcgagat gactgaccca ctggccaaca acaacgtcaa cgacctctc ctggacatgg 31440  
acggccggc ctcggagcag cgactcgccc aacttcgcat tcgccagcag caggagagag 31500  
ccgtcaagga gctgcaggac ggcatagcca tccaccagtg caagaaaggc atcttctgcc 31560  
tggtgaaaca ggccaagatc tcctacgagg tcaccagac cgaccatcgc ctctcctacg 31620  
agctcctgca gcagcgccag aagttcacct gcttggtcgg agtcaacccc atcgtcatca 31680  
cccagcagtc gggcgatacc aaggggtgca tccactgctc ctgcgactcc ccgactgcg 31740  
tccacactct gatcaagacc ctctcgggcc tcgcgcacct cctccccatg aactaatcac 31800  
ccccctatcc agtgaataaa agatcatatt gatgatttga gtttaataaa aataaagaat 31860  
cacttacttg aaatctgata ccaggctctc gtccatgttt tctgccaaca ccacttcaact 31920  
ccccctcttc cagctctggt actgcaggcc ccggcgggct gcaaaacttc tccacaccct 31980  
gaaggggatg tcaaattcct cctgtccctc aatcttcatt ttatcttcta tcagatgtcc 32040  
aaaaagcgcg tccgggtgga tgatgacttc gaccccgctc acccctacga tcgagacaac 32100  
gcaccgaccg tgcccttcat caaccccc ttcgtctctt cagatggatt ccaagagaag 32160  
ccccggggg tgctgtccct gcgtctggcc gatcccgtea ccaccaagaa cggggaaatc 32220  
accctcaagc tgggagatgg ggtggacctc gactcctcgg gaaaactcat ctccaacacg 32280  
gccaccaag cgcgccccc tctcagtttt tccaacaaca ccatttccct taacatggat 32340  
accctttttt acaacaacaa tggaaagtta ggcatgaaag tcaactgctcc actgaagata 32400  
ctagacacag acttgctaaa aacacttggt gtagcttatg gacaaggttt aggaacaaac 32460  
accactgggt cccttgttgc ccaactagca tccccacttg cttttgatag caatagcaaa 32520  
attgccctta atttaggcaa tggaccattg aaagtggatg caaatagact gaacatcaat 32580  
tgcaatagag gactctatgt tactaccaca aaagatgcac tggaaagcaa tataagttgg 32640  
gctaattgta tgacatttat aggaaatgcc atgggtgtca atattgatac acaaaaaggc 32700  
ttgcaatttg gcaccactag taccgtcgca gatgttaaaa acgcttaacc catacaaatc 32760  
aaacttggag ctggctctac atttgacagc acaggtgcaa ttggtgcatg gaacaaagat 32820  
gatgacaagc ttacactatg gaccacagcc gaccctctc caaattgtca catatattct 32880  
gaaaaggatg ctaagcttac actttgcttg acaaagtgtg gcagtcagat tctgggcact 32940  
gtttccctca tagctgttga tactggcagt ttaaatccca taacaggaac agtaaccact 33000  
gctcttctct cacttaaat ctagcaaat ggagttttgc aaagcagctc aacactagac 33060  
tcagactatt ggaatttcag acagggagat gttacacctg ctgaagccta tactaatgct 33120

---

-continued

---

ataggtttca tgcccaatct aaaagcatac cctaaaaaca caagtggagc tgcaaaaagt 33180  
cacattgttg ggaaagtga cctacatggg gatacaggca aacctgga cctcattatt 33240  
actttcaatg aaacaagtga tgaatcttg acttactgta ttaactttca atggcagtgg 33300  
ggggctgac aatataaaaa tgaaacactt gccgtcagtt cattcacctt ttctatatt 33360  
gctaaagaat aaacccccact ctgtacccca tctctgtcta tggaaaaaac tctgaaacac 33420  
aaaaataaat aaagtccaag tgttttattg attcaacagt tttacaggat tcgagcagtt 33480  
atttttcctc caccctoccca ggacatggaa tacaccaccc tctccccccg cacagccttg 33540  
aacatctgaa tgccattggt gatggacatg cttttggtct ccacgttcca cacagtttca 33600  
gagcgagcca gtctcgggtc ggtcagggag atgaaaccct ccgggcactc ccgcactctgc 33660  
acctcacagc tcaacagctg aggattgtcc tcggtggtcg ggatcacggg tatctggaag 33720  
aagcagaaga gcggcgttg gaatcatagt ccgcgaacgg gatcggccgg tgggtctgca 33780  
tcaggccccg cagcagctgc tgtcgcgcc gctccgtcaa gctgctgctc aggggtccg 33840  
ggtccagga ctcccctcag atgatgcca cggccctcag catcagctgt ctggtgcggc 33900  
ggggcagca gcgcagctg atctcgtca ggtcgtgca gtacgtgcaa cacaggacca 33960  
ccaggtggt caacagtcca tagttcaaca cgctccagcc gaaactcatc gcgggaagga 34020  
tgctaccac gtggccgtc taccagatcc tcaggtaaat caagtggcg cccctccaga 34080  
acacgctgcc catgtacatg atctccttg gcatgtggcg gttaccacc tcccgttacc 34140  
acatcacct ctggttgaac atgcagccc ggatgatcct gcggaaccac agggccagca 34200  
ccgccccgc cgcctcagc cgaagagacc ccgggtccc acaatggca tggaggacc 34260  
accgctcga cccgtggatc atctgggagc tgaacaagtc tatgttgca cagcacaggc 34320  
atatgctcat gcatctctc agcactctca gctcctcggg ggtcaaaacc atatcccagg 34380  
gcacgggaa ctcttcagc acagcgaacc ccgcagaaca gggcaatcct cgcacataac 34440  
ttacattgtg catggacagg gtatcgcaat caggcagcac cgggtgatcc tccaccagag 34500  
aagcgcgggt ctcggtctcc tcacagctg gtaagggggc cggccgatac ggggtatggc 34560  
gggacgggc tgatcgtgt cgcgaccgtg ttatgatgca gttgctttcg gacattttcg 34620  
tacttgctg agcagaacct ggtccggcg ctgcacacc atcgcggcg gcggtcccg 34680  
cgcttgaac gctcgtggt gaagttgtaa aacagccact ctctcagacc gtgcagcaga 34740  
tctagggcct caggagtgat gaagatccca tcatgcctga tggctctaata cacatcgacc 34800  
accgtggaat gggccagacc cagccagatg atgcaatctt gttgggttc ggtgacggcg 34860  
ggggaggaa gaacaggaag aacctgatt aacttttaata ccaaacggtc tcggagcact 34920  
tcaaatgaa gatcgcggag atggcacctc tcgccccgc tgtgttggtg gaaaataaca 34980  
gccaggtcaa aggtgatagc gttctogaga tgttccacgg tggcttccag caaagcctcc 35040  
acgcgcacat ccagaacaa gacaatagc aaagcgggag ggttctctaa ttctcaatc 35100  
atcatgttac actcctgcac catcccaga taattttcat tttccagcc ttgaatgatt 35160  
cgaactagtt cctgaggtaa atccaagcca gccatgataa agagctcgcg cagagcggcc 35220  
tccaccgca ttcttaagca caccctcata attccaagat attctgctcc tggttcacct 35280  
gcagcagatt gacaagcga atatcaaat ctctgcggc atccctaagc tcctccctca 35340  
gcaataactg taagtactct tcatatcct ctccgaaatt tttagccata ggaccaccag 35400

-continued

---

```

gaataagatt agggcaagcc acagtacaga taaaccgaag tcctcccag tgagcattgc 35460
caaatgcaag actgctataa gcatgctggc tagaccgggt gatatactcc agataactgg 35520
acagaaaatc gccagggcaa tttttaagaa aatcaacaaa agaaaaatcc tccaggtgca 35580
cgtttagagc ctcggaaca acgatggagt aaatgcaagc ggtgcttcc agcatggtta 35640
gtagctgat ctgtagaaaa aaacaaaaat gaacattaaa ccatgctagc ctggcgaaca 35700
ggtgggtaaa tcgttctctc cagcaccagg caggccacgg ggtctccggc acgaccctcg 35760
taaaaattgt cgctatgatt gaaaaccatc acagagagac gttcccggtg gccggcgtga 35820
atgattcgac aagatgaata ccccccgga acattggcgt ccgagagtga aaaaaagcgc 35880
ccaaggaagc aataaggcac tacaatgctc agtctcaagt ccagcaaagc gatgccatgc 35940
ggatgaagca caaaattctc aggtgctgac aaaatgtaat tactcccctc ctgcacaggc 36000
agcaaagccc ccgatccctc caggtacaca tacaagcctc cagcgtccat agcttaccga 36060
gcagcagcac acaacaggcg caagagtcag agaaaggctg agctctaacc tgtccacccg 36120
ctctctgctc aatataatagc ccagatctac actgacgtaa aggccaaagt ctaaaaatac 36180
ccgcaaaata atcacacacg cccagcacac gccagaaac cggtgacaca ctcaaaaaaa 36240
tacgcgcact tcctcaaagc cccaaactgc cgtcatttcc gggttcccac gctacgtcat 36300
caaaattcga ctttcaaatt ccgctgacgg ttaaaaacgt cgcccgcgcc gccccaacg 36360
gtcgcgctc ccgagccaa tcaccgccc gcaccccaa attcaatac ctcaattgca 36420
tattaacgcg caccaaaagt ttgaggtata ttattgatga tg 36462

```

&lt;210&gt; SEQ ID NO 2

&lt;211&gt; LENGTH: 36604

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: chimpanzee adenovirus serotype Pan6

&lt;400&gt; SEQUENCE: 2

```

catcatcaat aatatacctc aaacttttgg tgcgcgttaa tatgcaaatg agctgtttga 60
atgtggggag ggaggaaggt gattggctgc gggagcggcg accgttaggg gcggggcggg 120
tgacgttttg atgacgtggc tatgagcggg agccggtttg caagttctcg tgggaaaagt 180
gacgtcaaac gaggtgtggt ttgaacacgg aaatactcaa ttttcccgcg ctctctgaca 240
ggaaatgagg tgtttctggg cggatgcaag tgaaaacggg ccattttcgc gcgaaaactg 300
aatgaggaag tgaaaatctg agtaatttgc cgtttatggc agggaggagt atttgcgag 360
ggcgcagtag actttgaccg attacgtggg ggtttcgatt accgtatttt tcaactaaat 420
ttccgcgtac ggtgtcaaag tccggtgttt ttacgtaggc gtcagctgat cgcagggta 480
tttaaacctg cgctctctag tcaagaggcc actcttgagt gccagcagat agagttttct 540
cctccgcgcc gcgagtcaga tctacacttt gaaagatgag gcacctgaga gacctgccc 600
gtaatgtttt cctggctact gggaaacgaga ttctggaatt ggtggtggac gccatgatgg 660
gtgacgaccc tccagagccc cctaccccat ttgagggccc ttcgctgtac gattttgtatg 720
atctggaggt ggatgtgccc gagagcgacc ctaacgagga ggcggtgaat gattttgtta 780
gcgatgccgc gctgctggct gccgagcagg ctaatacggc ctctggctca gacagcgatt 840
cctctctcca taccgcgaga cccggcagag gtgagaaaa gatccccgag cttaaagggg 900
aagagctcga cctgcgctgc tatgaggaat gottgcctcc gagcagatgat gaggaggacg 960

```

-continued

---

|            |             |            |             |             |            |      |
|------------|-------------|------------|-------------|-------------|------------|------|
| aggaggcgat | tcgagctgcg  | gtgaaccagg | gagtgaaaac  | tgcgggcgag  | agctttagcc | 1020 |
| tggactgtcc | tactctgccc  | ggacacggct | gtaagtcttg  | tgaatttcat  | cgcatgaata | 1080 |
| ctggagataa | gaatgtgatg  | tgtgcocctg | gctatatgag  | agcttacaac  | cattgtgttt | 1140 |
| acagtaagtg | tgattaactt  | tagttgggaa | ggcagagggt  | gactgggtgc  | tgactggttt | 1200 |
| atztatgtat | atgttttttt  | atgtgtaggt | cccgtctctg  | acgtagatga  | gacccccact | 1260 |
| tcagagtgca | tttcatcacc  | cccagaaatt | ggcgaggaac  | cgcccgaaga  | tattattcat | 1320 |
| agaccagtgt | cagtgtagagt | caccggggcg | agagcagctg  | tggagagttt  | ggatgacttg | 1380 |
| ctacagggtg | ggatgaacc   | tttgacttg  | tgtaccggga  | aacggccag   | gcactaagtg | 1440 |
| ccacacatgt | gtgtttactt  | aagggtgatg | cagtatttat  | aggggtgtga  | gtgcaataaa | 1500 |
| atccgtgttg | actttaagtg  | cgtgttttat | gactcagggg  | tggggactgt  | gggtatataa | 1560 |
| gcaggtgcag | acctgtgtgg  | tcagttcaga | gcaggactca  | tggagatctg  | gactgtcttg | 1620 |
| gaagactttc | accagactag  | acagttgcta | gagaactcat  | cggagggagt  | ctcttacctg | 1680 |
| tggagattct | gcttcgggtg  | gcctctagct | aagctagtct  | atagggcaa   | acaggattat | 1740 |
| aaggaacaat | ttgaggatat  | tttgagagag | tgtccttgta  | tttttgactc  | tctcaacttg | 1800 |
| ggccatcagt | ctcactttaa  | ccagagtatt | ctgagagccc  | ttgacttttc  | tactcctggc | 1860 |
| agaactaccg | ccgcggtagc  | cttttttgcc | tttattcttg  | acaaatggag  | tcaagaacc  | 1920 |
| catttcagca | gggattaccg  | tctggactgc | ttagcagtag  | ctttgtggag  | aacatggagg | 1980 |
| tgccagcgcc | tgaatgcaat  | ctccggctac | ttgccagtac  | agccggtaga  | cacgctgagg | 2040 |
| atcctgagtc | tccagtcacc  | ccaggaacac | caacgcccgc  | agcagccgca  | gcaggagcag | 2100 |
| cagcaagagg | aggaccgaga  | agagaaccgg | agagccggtc  | tggaccctcc  | ggtggcggag | 2160 |
| gaggaggagt | agctgacttg  | tttcccgagc | tgcgcccggg  | gctgactagg  | tcttccagtg | 2220 |
| gacgggagag | ggggattaag  | cgggagaggc | atgaggagac  | tagccacaga  | actgaactga | 2280 |
| ctgtcagctc | gatgagccgc  | aggcggccag | aatcgggtgtg | gtggcatgag  | gtgcagtcgc | 2340 |
| aggggataga | tgaggtctcg  | gtgatgcatg | agaaatattc  | cctagaacaa  | gtcaagactt | 2400 |
| gttggttgga | gcccagggat  | gattgggagg | tagccatcag  | gaattatgcc  | aagctggctc | 2460 |
| tgaagccaga | caagaagtac  | aagattacca | aactgattaa  | tatcagaaat  | tctgtctaca | 2520 |
| tttcagggaa | tggggccgag  | gtggagatca | gtaccagga   | gagggtggcc  | ttcagatggt | 2580 |
| gtatgatgaa | tatgtaccgg  | ggggtggtgg | gcatggaggg  | agtcaccttt  | atgaacacga | 2640 |
| ggttcagggg | tgatgggtat  | aatggggtgg | tctttatggc  | caacaccaag  | ctgacagtgc | 2700 |
| acggatgctc | cttctttggc  | ttcaataaca | tgtgcatcga  | ggcctggggc  | agtgtttcag | 2760 |
| tgaggggatg | cagcttttca  | gccaactgga | tgggggtcgt  | gggcagaacc  | aagagcaagg | 2820 |
| tgtcagtgaa | gaaatgcctg  | ttcgagaggt | gccacctggg  | ggtgatgagc  | gagggcgaag | 2880 |
| ccaaagtcaa | acactgcgcc  | tctaccgaga | cgggctgctt  | tgtgctgac   | aagggcaatg | 2940 |
| cccaagtcaa | gcataacatg  | atctgtgggg | cctcggatga  | gcgcggtac   | cagatgctga | 3000 |
| cctgcgcccg | tgggaacagc  | catatgctgg | ccaccgtgca  | tgtggcctcg  | cacccccgca | 3060 |
| agacatggcc | cgagttcgag  | cacaacgtca | tgaccgctg   | caatgtgcac  | ctgggctccc | 3120 |
| gccgaggcat | gttcatgccc  | taccagtgca | acatgcaatt  | tgtgaagggtg | ctgctggagc | 3180 |
| ccgatgccat | gtccagagtg  | agcctgacgg | gggtgtttga  | catgaatgtg  | gagctgtgga | 3240 |

-continued

---

|            |            |             |            |             |             |      |
|------------|------------|-------------|------------|-------------|-------------|------|
| aaattctgag | atatgatgaa | tccaagacca  | ggtgccgggc | ctgcgaatgc  | ggaggcaagc  | 3300 |
| acgccagget | tcagcccgtg | tgtgtggagg  | tgacggagga | cctgcgaccc  | gatcatttgg  | 3360 |
| tgtgtcctg  | caacgggacg | gagttcggct  | ccagcgggga | agaatctgac  | tagagtgagt  | 3420 |
| agtgtttggg | gctgggtgtg | agcctgcatg  | aggggcagaa | tgactaaaat  | ctgtggtttt  | 3480 |
| ctgtgtgttg | cagcagcatg | agcgggaagc  | cctcctttga | gggaggggta  | ttcagccctt  | 3540 |
| atctgacggg | gcgtctcccc | tcctggggcg  | gagtgcgca  | gaatgtgatg  | ggatccacgg  | 3600 |
| tgacggcgcg | gcccgtgcag | cccgcgaact  | cttcaaccct | gacctacgcg  | accctgagct  | 3660 |
| cctcgtccgt | ggacgcagct | gcccgcgag   | ctgctgcttc | cgccgccagc  | gccgtgcccg  | 3720 |
| gaatggccct | ggcgccggc  | tactacagct  | ctctggtggc | caactcgagt  | tccaccaata  | 3780 |
| atcccgcag  | cctgaacgag | gagaagctgc  | tgctgctgat | ggcccagctc  | gaggccctga  | 3840 |
| cccagcgcct | ggcgagctg  | accagcagc   | tggtcagct  | gcaggcggag  | acgcggggcg  | 3900 |
| cggttgccac | ggtgaaaacc | aaataaaaa   | tgaatcaata | aataaacgga  | gacggttggt  | 3960 |
| gattttaaca | cagagtcttg | aatctttatt  | tgatttttcg | cgccggtag   | gccctggacc  | 4020 |
| accggtctcg | atcattgagc | accgggtgga  | tcttttcag  | gacccggtag  | aggtgggctt  | 4080 |
| ggatgttgag | gtacatgggc | atgagcccgt  | cccgggggtg | gaggtagctc  | cattgcaggg  | 4140 |
| cctcgtgctc | gggatgggtg | ttgtaaatca  | cccagtcata | gcagggggcg  | aggcgtggt   | 4200 |
| gctgcacgat | gtccttgagg | aggagactga  | tggccacggg | cagccccttg  | gtgtaggtgt  | 4260 |
| tgacgaacct | gttgagctgg | gagggatgca  | tgcgggggga | gatgagatgc  | atcttggcct  | 4320 |
| ggatcttgag | attggcgatg | ttcccgccca  | gatcccgcg  | ggggttcattg | ttgtgcagga  | 4380 |
| ccaccagcac | ggtgtatccg | gtgcacttgg  | ggaatttgtc | atgcaacttg  | gaagggaaag  | 4440 |
| cgtagaaaga | tttgagagc  | cccttgtagc  | cgcccagggt | ttccatgcac  | tcacccatga  | 4500 |
| tgatggcgat | ggcccgtgg  | gcggcggcct  | gggcaaagac | gtttcggggg  | tcggacacat  | 4560 |
| cgtagttgtg | gtcctgggtg | agctcgtcat  | aggccatttt | aatgaatttg  | ggcgggaggg  | 4620 |
| tgcccgactg | gggacgaag  | gtgccctcga  | tcccgggggc | gtagttgcc   | tcgcagatct  | 4680 |
| gcatctccca | ggccttgagc | tcggaggggg  | ggatcatgct | cacctgcggg  | gcgatgaaaa  | 4740 |
| aaacggttcc | cgggcgggg  | gagatgagct  | gggcccgaag | caggttcgg   | agcagctggg  | 4800 |
| acttgccgca | accggtgggg | ccgtagatga  | ccccgatgac | cggtgcagc   | tggtagtga   | 4860 |
| gggagagaca | gctgcctgcc | tcgcggagga  | ggggggccac | ctcgttcac   | atctcgcgca  | 4920 |
| catgcatggt | ctcgcgcacg | agttccgccca | ggaggcgtc  | gccccccagc  | gagaggagct  | 4980 |
| cttgacgca  | ggcgaagttt | ttcagcggct  | tgagtccgtc | ggccatgggc  | atthttggaga | 5040 |
| gggtctgttg | caagagttcc | agacggtccc  | agagctcgg  | gatgtgctct  | agggcatctc  | 5100 |
| gatccagcag | acctcctcgt | ttcgcgggtt  | ggggcgactg | cgggagtagg  | gcaccagggc  | 5160 |
| atgggcgtcc | agcgaggcca | gggtccggct  | cttccagggc | cgaggggtcc  | gcgtcagcgt  | 5220 |
| ggtctccgtc | acggtgaagg | ggtgcgcgcc  | gggctggggc | cttgcgaggg  | tcgccttcag  | 5280 |
| gctcatccgg | ctggtcgaga | accgctccc   | gtcggcgccc | tcgcgctcgg  | ccaggtagca  | 5340 |
| attgagcatg | agttcgtagt | tgagcgcctc  | ggccgcgtgg | cccttgccgc  | ggagcttacc  | 5400 |
| tttgaagtg  | tgccgcgaga | cgggacagag  | gagggacttg | agggcgtaga  | gcttgggggc  | 5460 |
| gaggaagacg | gactcggggg | cgtaggcgtc  | cgccgcgag  | ctggcgcgaga | cgtctcgcga  | 5520 |

-continued

---

|   |      |
|---|------|
| ctccacgagc caggtgaggt cggggcgggt ggggtcaaaa acgaggttcc ctccgtgctt | 5580 |
| ttgatgcgt ttcttacctc tggctccat gagctcgtgt ccccgctggg tgacaaagag   | 5640 |
| gctgtccgtg tccccgtaga ccgactttat gggccggtcc tcgagcgggg tgccgcggtc | 5700 |
| ctcgtcgtag aggaaccccg cccactccga gacgaaggcc cgggtccagg ccagcacgaa | 5760 |
| ggagggccacg tgggaggggt agcggtcggt gtccaccagc gggccacct tctccagggt | 5820 |
| atgcaagcac atgtcccct cgtccacatc caggaagggt attggctgt aagtgtaggc   | 5880 |
| cacgtgaccg ggggtcccgg ccgggggggt ataaaagggg gcgggcccct gctcgtcctc | 5940 |
| actgtcttcc ggatcgtgt ccaggagcgc cagctgttg ggtaggatt ccctctcgaa    | 6000 |
| ggcgggcatg acctcggcac tcaggttgtc agtttctaga aacgaggagg atttgatatt | 6060 |
| gacggtgccc ttggagagc ctttcatgag cccctcgtcc atttggcag aaaagacgat   | 6120 |
| ctttttgttgc tcgagcttg tggcgaagga gccgtagagg gcgttgaga gcagcttggc  | 6180 |
| gatggagcgc atggctcgtt tcttttcctt gtcggcgcgc tccttgccgg cgatgtgag  | 6240 |
| ctgcacgtac tcgcgcgcca cgcacttcca ttcggggaag acggtggtga gctcgtcggg | 6300 |
| cacgattctg acccgcaccg cgcggttggt cagggtagt aggtccacgc tggtgggccac | 6360 |
| ctcgcgcgc aggggctcgt tggccacga gaggcgccc cccttgccgc agcagaaggg    | 6420 |
| gggcagcggg tccagcatga gctcgtcggg ggggtcggcg tccacggtga agatgccggg | 6480 |
| caggagctcg gggcgaagt agctgatgca ggtgccaga ttgtccagcg ccgcttgcca   | 6540 |
| gtcgcgcacg gccagcgcgc gctcgtaggg gctgaggggc gtgcccagg gcatggggtg  | 6600 |
| cgtgagcgcg gaggcgtaca tgcgcagat gtcgtagacg tagaggggct cctcagggac  | 6660 |
| gccgatgtag gtgggttagc agcgcctccc gcggtgctg gcgcgcacgt agtcgtacag  | 6720 |
| ctcgtgcgag ggcgcgagga gcccctgccc gaggttgag cgttgccgct tttcggcgcg  | 6780 |
| gtagacgac tggcgaaga tggcgtggga gttggaggag atggtgggccc tttggaagat  | 6840 |
| gttgaagtgg gcgtgggga gcccgaccga gtcccctgat aagtgggctg aggagctctg  | 6900 |
| cagcttgccg acgagctcgg cggtagcag gacgtccagg gcgcagtagt cgagggtctc  | 6960 |
| ttggatgatg tcatacttga gctggccctt ctgcttccac agctcgcggt tgagaaggaa | 7020 |
| ctcttcgagg tccttccagt actcttcgag ggggaaccgg tcctgatcgg cacggtaaga | 7080 |
| gcccaccatg tagaactggt tgacggcctt gtaggcgag cagcccttct ccacggggag  | 7140 |
| ggcgtaagct tgcgcggcct tgcgcaggga ggtgtgggtg agggcgaagg tgcgcgcac  | 7200 |
| catgaccttg aggaactggt gcttgaagtc gaggtcgtc cagccgcct gctccagag    | 7260 |
| ttggaagtcc gtgcgcttct ttagggcgg gttaggcaaa gcgaaagtaa catcgtttaa  | 7320 |
| gaggatcttg cccgcgagg gcatgaagtt gcgagtagt cggaaaggct ggggcacctc   | 7380 |
| ggcccggtt ttgatgacct gggcggcgag gacgatctc tcgaagcgt tgatgttgtg    | 7440 |
| cccgacgatg tagagttcca cgaatcgcgg gcggcccttg acgtgggca gcttcttgag  | 7500 |
| ctcgtcgtag gtgagctcgg cggggtcgt gagccctgc tgcctcaggg cccagtcggc   | 7560 |
| gacgtggggg ttggcgtga ggaaggaaat ccagagatcc acggccaggg cggctctgcaa | 7620 |
| gcggtcccgg tactgacgga actgttgccc cacggccatt tttcggggg tgacgcagta  | 7680 |
| gaaggtcgg gggctcggct gccagcggtc ccacttgagc tggagggcga ggtcgtgggc  | 7740 |
| gagctcgcag agcggcgggt ccccgagag tttcatgacc agcatgaag ggaagagctg   | 7800 |



---

-continued

---

cttgccgaag gaccccatcc aggtgtaggt ttccacatcg taggtgagga agagcctttc 7860  
gggtcgagga tgcgagccga tggggaagaa ctggatctcc tgcaccacgt tggaggaatg 7920  
gctgttgatg tgatggaagt agaaatgccg acggcgcgcc gagcactcgt gcttgtgttt 7980  
atacaagcgt ccgcatgtct cgcaacgctg cacgggatgc acgtgctgca cgagctgtac 8040  
ctgggttctt ttggcgagga atttcagtgg gcagtggagc gctggcggtc gcatctcgtg 8100  
ctgtactacg tcttgccat cggcgtggcc atcgtctgcc tcgatggtgg tcatgctgac 8160  
gagccccgcg gggaggcagg tccagacctc ggctcggacg ggtcggagag cgaggacgag 8220  
ggcgcgcagg ccgagctgt ccagggtcct gagacgctgc ggagtcaggt cagtgggcag 8280  
cggcgcgcg cggttgactt gcaggagctt ttccaggcg cgcgggaggt ccagatggta 8340  
cttgatctcc acggcgccgt tggtggtac gtccacggct tgcagggtgc cgtgccctc 8400  
ggggccacc accgtgcccc gtttctctt gggcgctgct tccatgtcgg tcagaagcgg 8460  
cggcgaggac gcgcgccggg cggcaggggc ggctcggggc ccgaggcag gggcggcagg 8520  
ggacgctcgg cgcgcgcgc gggcagggtc tgggtactcg cccggagaag actggcgtga 8580  
gcgacgacgc gacgggtgac gtcctggatc tgacgcctct ggggaaggc cacgggacct 8640  
gtgagtttga acctgaaaga gaggctgaca gaatcaatct cggtatcgtt gacggcggcc 8700  
tgccgcagga tctcttgac gtcgcccag ttgtcctggt aggcgatctc ggtcatgaac 8760  
tgctcgatct cctcctcctg aaggctcctg cggccggcgc gctcgcaggt ggcgcgag 8820  
tcgttgaga tgcggccat gactgagcag aaggcgttca tgcggcctc gttccagacg 8880  
cggctgtaga ccacgctcc gtcggggtcg cgcgcgcga tgaccacctg ggcgaggtt 8940  
agctcgacgt ggcgcgtgaa gaccgcgtag ttgcagaggc gctggtagag gtagttgagc 9000  
gtggtggcga tgtgctcgtg gacgaagaag tacatgatcc agcggcgag cgcatctc 9060  
ctgacgtcgc ccaggcttc caagcgttc atggcctcgt agaagtccac ggcgaagtt 9120  
aaaaactggg agttgcgcg cgagaaggtc aactcctcct ccagaagacg gatgagctc 9180  
gcgatggtg cgcgcacctc gcgctogaag gcccggggg gctcctcttc catctcctc 9240  
tctcctcct ccactaacat ctctctact tctcctcag gaggcgttg cgggggagg 9300  
gccctgcgtc gccggcgcg cacgggcaga cggtcgatga agcgtcgtg ggtctcccc 9360  
cgccgcgac gcatggtctc ggtgacgcg cgcctcctc cgcgggccg cagcatgaag 9420  
acgccgcgc gcatctccag gtggcgcgc ggggggtctc cgttgggcag ggagagggc 9480  
ctgacgatgc atcttatcaa ttgaccgta gggactccgc gcaaggacct gagcgtctc 9540  
agatccacgg gatccgaaaa ccgctgaacg aaggcttca gccagtcgca gtcgcaagg 9600  
aggctgagcc cggttctctg tctctcgggt atttggtcgg gaggcggcg ggcgatgct 9660  
ctggtgatga agttgaagta ggcggtcctg agacggcgga tgggtggcag gacaccagg 9720  
tccttgggcc cggcttgctg gatgcgcaga cggtcggcca tccccaggc gtggtcctga 9780  
cacctggcga ggtccttga gtagtctgc atgagccgt ccacgggac ctcctcctc 9840  
cccgcgcgc cgtgcatgcg cgtgagccc aaccgcgct cggcgtggac gagcgccag 9900  
tcggcgacga cgcgctcgt gaggatggcc tgctggatct gggtaggggt ggtctggaag 9960  
tcgtcaagt cgacgaagc gtggtaggct ccggtgttga tgggttagga gcagttggc 10020  
atgacggacc agttgacggt ctggtggccg ggtcgacga gctcgtgta cttgagggc 10080

---

-continued

---

gagtaggcgc gcgtgtcgaa gatgtagtgc ttgcaggcgc gcacgaggta ctggtatccg 10140  
acgaggaagt gcggcggcgg ctggcggtag agcggccatc gctcgggtggc gggggcgcgc 10200  
ggcgcgaggt cctcagcat gaggcggtgg tagccgtaga tgtacctgga catccaggtg 10260  
atgccggcgg cgggtggtgga ggcgcgcggg aactcgcgga cgcggttcca gatggtgcgc 10320  
agcggcagga agtagttcat ggtggcccg gtctggcccg tgaggcgcgc gcagtcgtgg 10380  
atgctctaga catacgggca aaaacgaaag cggtcagcgg ctcgactccg tggcctggag 10440  
gctaagcgaa cgggttgggc tgcgcgtgta ccccggttcg aatctcgaat caggctggag 10500  
ccgcagctaa cgtggtactg gcactcccgt ctcgacccaa gcctgctaac gaaacctcca 10560  
ggatacggag gcgggtcgtt ttttggcctt ggtcgcctgg catgaaaaac tagtaagcgc 10620  
ggaaagcggc cccccgcgat ggctcgcctc cgtagtctgg agaagaatc gccagggttg 10680  
cgttgccgtg tgcctccgtt cgagcctcag cgtcggcgc cggccggatt ccgcggctaa 10740  
cgtggcgtg gctgcccctg cgtttccaag accccttagc cagccgactt ctccagttac 10800  
ggagcagcc cctctttttt tttcttgtgt ttttgccaga tgcacccgt actgcccag 10860  
atgcccccc accctccacc acaaccgccc ctaccgcagc agcagcaaca gccggcgtt 10920  
ctgccccgc cccagcagca gccagccact accgcggcgg ccgccgtgag cggagccggc 10980  
gttcagtatg acctggcctt ggaagagggc gaggggctgg cgcggctggg ggcgtcgtcg 11040  
ccggagcggc acccgcgcgt gcagatgaaa agggacgctc gcgaggccta cgtgcccaag 11100  
cagaacctgt tcagagacag gagcggcag gagcccagg agatgcgcgc ctcccgttc 11160  
cacgcggggc gggagctgc gcgcggcctg gaccgaaagc ggtgctgag ggacgaggat 11220  
ttcaggcggc acgagctgac ggggatcagc cccgcgcgcg cgcacgtggc cgcggccaac 11280  
ctggtcacgg cgtacgagca gaccgtgaag gaggagagca acttccaaaa atccttcaac 11340  
aaccacgtgc gcacgtgat cgcgcgcgag gaggtgacc tggcctgat gcacctgtg 11400  
gacctgctg aggccatcgt gcagaacccc acgagcaagc cgtgacggc gcagctgttt 11460  
ctggtggtgc agcacagtgc ggacaacgag acgttcaggg aggcgctgct gaataacc 11520  
gagcccagg gccgctggct cctggacctg gtgaacattt tgcagagcat cgtggtgcag 11580  
gagcgcgggc tgcctgtgc cgagaagctg gcggccatca acttctcggg gctgagtctg 11640  
ggcaagtact acgctaggaa gatctacaag acccgtacg tgcccataga caaggagggtg 11700  
aagatcgac gggtttatcat gcgcatgacc ctgaaagtgc tgaccctgag cgacgatctg 11760  
ggggtgtacc gcaacgacag gatgcaccgc gcggtgagc ccagcccgc gcgcgagctg 11820  
agcgaccagg agctgatgca cagcctgag cgggccctga ccggggcgcg gaccgagggg 11880  
gagagctact ttgacatggg cgcggacctg cgtggcagc ccagcccgcg ggccttggaa 11940  
gctgcccggc gttcccccta cgtggaggag gtggacgatg aggaggagga gggcaggtac 12000  
ctggaagact gatggcgcga ccgtattttt gctagatgca gcaacagcca ccgcccccgc 12060  
ctctgatcc cgcgatgcgg gcggcgtgc agagccagcc gtcggcatt aactcctcg 12120  
acgattggac ccaggccatg caacgcacga tggcgtgac gaccgcgaat cccgaagcct 12180  
ttagacagca gcctcaggcc aaccggctct cggccatcct ggaggccgtg gtgccctcgc 12240  
gctcgaaccc cacgcacgag aaggtgctgg ccatcgtgaa cgcgctggtg gagaacaagg 12300  
ccatccgcgg tgacgaggcc gggctggtgt acaacgcgct gctggagcgc gtggcccgc 12360

---

-continued

---

acaacagcac caacgtgcag acgaacctgg accgcatggt gaccgacgtg cgcgagggcgg 12420  
tgtgcgagcg cgagcgggttc caccgcgagt cgaacctggg ctccatggtg gcgctgaacg 12480  
ccttcctgag cagcagcccc gccaacgtgc cccggggcca ggaggactac accaacttca 12540  
tcagcgcgct gcggctgatg gtggccgagg tgccccagag cgagggtgtac cagtcggggc 12600  
cggactactt cttccagacc agtcgccagg gcttgacagc cgtgaacctg agccaggctt 12660  
tcaagaactt gcagggactg tggggcgtgc agggccccgt cggggaccgc gcgacgggtg 12720  
cgagcctgct gacgcgcaac tcgcccctgc tgctgctgct ggtggcggcc ttcacggaca 12780  
gcggcagcgt gacccgagc tcgtacctgg gctacctgct taacctgtac cgcgagggcca 12840  
tcggacaggc gcacgtggac gagcagacct accaggagat caccacgtg agccgcgcgc 12900  
tgggccagga ggaccgggc aacctggagg ccacctgaa cttcctgctg accaaccggg 12960  
cgcagaagat cccgccccag tacgcgctga gcaccgagga ggagcgcac ctcgctacg 13020  
tgacgacag cgtggggctg ttcctgatgc aggagggggc cacgcccagc gcggcgcctc 13080  
acatgaccgc gcgcaacatg gagcccagca tgtacgccc caaccgccc ttcacata 13140  
agctgatgga ctacttgcat cggcggggc ccatgaactc ggactacttt accaacgcca 13200  
tcttgaacc gcactggctc ccgcccggc ggttctacac gggcgagtac gacatgccc 13260  
acccaacga cgggttcctg tgggacgagc tggacagcag cgtgttctc ccgctccag 13320  
gaaccaatgc cgtgtggaag aaagagggcg gggaccggc gccgtcctc gcgctgtccg 13380  
gtcgcgcggg tgctgcccgc gcggtgccc aggcccagc ccccttccc agcctgcct 13440  
tttcgctgaa cagcgtgccc agcagcagc tgggtcggct gacgcgacc cgcctgctg 13500  
gcgagggagg gtacctgaac gactccttgt tgaggcccga gcgcgagaag aacttcccc 13560  
ataacgggat agagagcctg gtggacaaga tgagccgctg gaagacgtac gcgcacgagc 13620  
acagggacga gccccgagct agcagcagc gcacctgag acgcccagc cagcagggc 13680  
agcggggact ggtgtgggac gatgaggatt ccgcccagc cagcagcgtg ttggacttgg 13740  
gtgggagtgg tgtaacccg ttcgctcacc tgcgcccgc tctcggggc ctgatgtaag 13800  
aatctgaaaa aataaaagac ggtactcacc aaggccatgg cgaccagcgt gcgttcttct 13860  
ctgtgtttt tagtagtatg atgagggcgc tgtacccgga ggtcctcct cctcgtacg 13920  
agagcgtgat gcagcagggc gtggcggcg cgatgcagcc cccgctggag gcgccttacg 13980  
tgccccgcg gtacctggcg cctacggagg ggcggaacag cattcgttac tcggagctgg 14040  
caccttcta cgataccacc cggttgtacc tggtgacaa caagtccgca gacatcgcct 14100  
cgctgaacta ccagaacgac cacagcaact tcctgaccac cgtggtgcag aacaacgatt 14160  
tcacccccac ggaggccagc acccagacca tcaactttga cgagcgcctc cgtggggcg 14220  
gccagctgaa aaccatcatg cacaccaaca tgcccacgt gaacgagttc atgtacagca 14280  
acaagtcaa gcgcggggg atggtctcgc gcaagacccc caacggggg gatgatgatt 14340  
atgatgtag tcaggacgag ctgacctac agtgggtgga gtttgagctg cccgagggca 14400  
acttctcgtg gacctgacc atcgatctga tgaacaacgc catcatcgac aactacttgg 14460  
cggtgggggc gcagaacggg gtgctggaga gcgacatcgc cgtgaagtcc gacacgcgca 14520  
acttccggct gggctgggac cccgtgacc agctgggtgat gccgggcgtg tacaccaacg 14580  
aggccttcca ccccgacatc gtcctgctgc ccgctgcgg cgtggacttc accgagagcc 14640

---

-continued

---

gcctcagcaa cctgctgggc atccgcaagc ggcagccctt ccaggagggc ttccagatcc 14700  
tgtacgagga cctggagggg ggcaacatcc ccgcgctctt ggatgtcgaa gcctacgaga 14760  
aaagcaagga ggatagcacc gcccgggcga ccgcagccgt ggccaccgcc tctaccgagg 14820  
tgcggggcga taatthttgct agcgtgctgg cagcgccga ggcggctgaa accgaaagta 14880  
agatagtcat ccagccgggtg gagaaggaca gcaaggacag gagctacaac gtgctcggg 14940  
acaagaaaa caccgcctac cgcagctggt acctggccta caactacggc gaccccgaga 15000  
agggcgctgc ctctctggag ctgctcacca cctcggacgt cacctcgggc gtggagcaag 15060  
tctactggtc gtgccccgac atgatgcaag acccggtcac cttccgctcc acgctcaag 15120  
ttagcaacta cccggtgggtg ggcgcgagc tcctgcccgt ctactccaag agctttctca 15180  
acgagcaggc cgtctactcg cagcagctgc gcgccttcc ctcgctcacg cacgtcttca 15240  
accgcttccc cgagaaccag atcctcgtcc gcccgcccgc gccaccatt accaccgtca 15300  
gtgaaaacgt tcctgctctc acagatcacg ggaccctgcc gctgcgcagc agtatccggg 15360  
gagtcacgag cgtgaccgtc actgacgcca gacgccgcac ctgcccctac gtctacaagg 15420  
ccctggggtg agtgcgcgag cgcgtctctc cgagccgcac cttctaaaaa atgtccattc 15480  
tcactctgcc cagtaataac accggttggg gcctgcgcgc gccagcaag atgtacggag 15540  
gcgctcgcca acgctccacg caacccccg tgcgctgctg cgggcacttc cgcgctccct 15600  
ggggcgccct caagggccgc gtgctctcgc gcaccaccgt cgacgacgtg atcgaccagg 15660  
tggtggccga cgcgcgcaac tacacgcccg ccgcgcgccc cgtctccacc gtggacgccg 15720  
tcactgacag cgtggtggcc gacgcgcgcc ggtacgcccg caccaagagc cggcgcgggc 15780  
gcatcggccc gcggcaccgg agcacccccg ccatgcgcgc ggcgcgagcc ttgctgcgca 15840  
gggccaaggc cacgggacgc agggccatgc tcaggggggc cagacgcgcg gcctccggca 15900  
gcagcagcgc cggcaggacc cgcagacgcg cggccacggc ggcggcgggc gccatcgcca 15960  
gcatgtcccc ccccgcgcgc ggcaacgtgt actgggtgct cgacgccgcc accggtgtgc 16020  
gcgtgcccgt gcgcaccgc ccccctcgca cttgaagatg ctgacttcgc gatggtgatg 16080  
tgtcccagcg gcgagaggga tgtccaagcg caaatacaag gaagagatgc tccaggtcat 16140  
cgcgcctgag atctacggcc ccgcgcgggc ggtgaaggag gaaagaaagc cccgaaact 16200  
gaagcgggtc aaaaaggaca aaaaggagga ggaagatgac ggactggtgg agtttctgctg 16260  
cgagttcgcc ccccgcgcgc gcgtgcagtg gcgcggggcg aaagtgaac cggctgctgcg 16320  
gcccgccacc acggtggtct tcacgcccgg cgagcgttcc ggctccgctt ccaagcgtc 16380  
ctacgacgag gtgtacgggg acgaggacat cctcagacag gcggtcgagc gtctgggcca 16440  
gtttgcgtac ggcaagcgca gccgccccgc gccctgaaa gaggaggcgg tgtccatccc 16500  
gctggaccac ggcaacccca cgcgagcct gaagccggtg acctgcagc aggtgctacc 16560  
gagcgcggcg ccgcgcgggg gcttcaagcg cgagggcggc gaggatctgt acccgaccat 16620  
gcagctgatg gtgcccgaag gccagaagct ggaggacgtg ctggagcaca tgaaggtgga 16680  
ccccgaggtg cagcccagag tcaaggtgct gccatcaag caggtggccc cgggcctggg 16740  
cgtgcagacc gtggacatca agatccccac ggagcccatg gaaacgcaga ccgagcccgt 16800  
gaagcccagc accagcacca tggaggtgca gacggatccc tggatgccc caccagcttc 16860  
caccagcact cgcggaagac gcaagtacgg cgcggccagc ctgctgatgc ccaactacgc 16920

---

-continued

---

gctgcatcct tccatcatcc ccacgccggg ctaccgccgc acgcgcttct accgcggcta 16980  
caccagcagc cgcgcccgca agaccaccac ccgccgccgt cgtcgcagcc gccgcagcag 17040  
caccgcgact tccgccttgg tgcggagagt gtatcgcagc gggcgcgagc ctctgaccct 17100  
gccgcgcgcy cgctaccacc cgagcatcgc catttaacta ccgcctccta cttgcagata 17160  
tggccctcac atgccgcctc cgcgtcccca ttacgggcta ccgaggaaga aagccgcgcc 17220  
gtagaaggct gacggggaac gggctgcgtc gccatcacca ccggcggcgg cgcgccatca 17280  
gcaagcgggt gggggggagg ttccctgccg cgctgatccc catcatcgcc gggcgcatcg 17340  
ggcgcatccc cggcatagct tccgtggcgg tgcagcctc tcagcggcac tgagacacaa 17400  
aaaagcatgg atttgaata aaaaaaaaaa tggactgacg ctccctgtcc tgtgatgtgt 17460  
gtttttagat ggaagacatc aattttctgt ccctggcacc gcgacacggc accgcggcct 17520  
ttatgggac ctggagcgac atcggcaaca gccaaactgaa cgggggcgcc ttcaattgga 17580  
gcagtctctg gagcgggctt aagaatttcg ggtccacgct caaacctat ggcaacaagg 17640  
cgtggaacag cagcacaggc caggcgtga gggaaaagct gaaagaacag aacttccagc 17700  
agaaggtggt tgatggcctg gcctcaggca tcaacggggt ggttgacctg gccaacccag 17760  
ccgtgcagaa acagatcaac agccgcctgg acgcggctcc gcccgcgggg tccgtggaga 17820  
tgccccaggt ggaggaggag ctgcctcccc tggacaagcg cggcgacaag cgaccgcgctc 17880  
ccgacgcgga ggagacgctg ctgacgcaca cggacgagcc gccccgtac gaggaggcgg 17940  
tgaaactggg cctgcccacc acgcggcccg tggcgcctct ggccaccgga gtgctgaaac 18000  
ccagcagcag ccagcccgcg accctggact tgcctccgcc tcgcccctcc acagtggcta 18060  
agcccctgcc gccggtggcc gtcgcgtcgc gcgcccccg aggcgcgccc caggcgaact 18120  
ggcagagcac tctgaacagc atcgtgggtc tgggagtgca gagtgtgaag cgcgcgctct 18180  
gtattaaaa gacactgtag cgcttaactt gcttgtctgt gtgtatatgt atgtccgctg 18240  
accagaagga ggagtgtgaa gaggcgcgtc gccgagttgc aagatggcca ccccatcgat 18300  
gctgccccag tggcggtaca tgcacatcgc cggacaggac gcttcggagt acctgagtcc 18360  
gggtctggtg cagtctgccc gcgccacaga cacctacttc agtctgggga acaagttag 18420  
gaacccacg gtggcgcaca cgcacgatgt gaccaccgac cgcagccagc ggctgacgct 18480  
gcgcttcgty cccgtggacc gcgaggacaa cacctactcg taaaaagtgc gctacacgct 18540  
ggcgtgggc gacaaccgcy tgcctggacat ggccagcacc tactttgaca tccgcggcgt 18600  
gctggaccgg gccctagct tcaaacctca ctctggcacc gcctacaaca gcctagctcc 18660  
caaggagct cccaattcca gccagtggga gcaagcaaaa acaggcaatg ggggaactat 18720  
ggaaacacac acatattggt tggcccctaat gggcggagag aatattacaa aagatggtct 18780  
tcaaattgga actgacgtta cagcgaatca gaataaacca atttatgccg acaaacatt 18840  
tcaaccagaa ccgcaagtag gagaagaaaa ttggcaagaa actgaaaact tttatggcgg 18900  
tagagctctt aaaaaagaca caaacatgaa accttgctat ggctcctatg ctagaccac 18960  
caatgaaaaa ggaggtcaag ctaaaactaa agttggagat gatggagttc caaccaaga 19020  
attgcacata gacctggctt tctttgatac tcccgggtggc accgtgaacg gtcaagacga 19080  
gtataaagca gacattgtca tgtataccga aaacacgtat ttggaaactc cagacacgca 19140  
tgtgtatata aaaccaggca aggatgatgc aagttctgaa attaacctgg ttcagcagtc 19200

---

-continued

---

tatgccaac agaccaact acattgggtt cagggacaac tttatcggtc ttatgtacta 19260  
caacagcact ggcaatatgg gtgtgcttgc tggtcaggcc tcccagctga atgctgtggt 19320  
tgatttgcaa gacagaaaca ccgagctgtc ctaccagctc ttgcttgact ctttgggtga 19380  
cagaaccggy tatttcagta tgtggaacca ggcggtggac agttatgacc ccgatgtgcy 19440  
catcatcgaa aacctaggty tggaggatga attgccaaac tattgcttcc ccttgagcgy 19500  
ctctggcact aacgccgat accaaggtyt gaaagtaaaa gatggtcaag atggtgatgt 19560  
tgagagtga tgggaaaatg acgatactgt tgcagctcga aatcaattat gtaaaggtaa 19620  
cattttcgcc atggagatta atctccagcy taacctgtgg agaagtttcc tctactcgaa 19680  
cgtggccctg taacctcccg actcctacaa gtacacgccc accaacgtca cgtgcccgc 19740  
caacaccaac acctacgatt acatgaatgg cagagtgaca cctccctcgc tggtagacgc 19800  
ctacctcaac atcggggcyg gctggtcgtt ggaccccatg gacaacgtca accccttcaa 19860  
ccaccaccgc aacgcgggcy tgcgctaccg ctccatgctc ctgggcaacg ggcgctacgt 19920  
gcccttccac atccaggtgc cccaaaagt tttcgccatc aagagcctcc tgcctcgtcc 19980  
cgggtcctac acctacgagt ggaacttccg caaggacgcy aacatgatcc tgcagagctc 20040  
cctaggcaac gacctgcyga cggacgggcy ctccatcgc ttcaccagca tcaacctcta 20100  
cgccaccttc ttcccatggy cgcacaacac cgcctccacg ctcgaggcca tgcgcycaa 20160  
cgacaccaac gaccagctct tcaacgacta cctctcggcy gccaacatgc tctaccccat 20220  
cccggccaac gccaccaacg tgcctatctc catcccctcy cgcactggg cgccttccg 20280  
cggatggtcc ttcaacgcyg tgaagaccgy cggagacccc tgcctcggct cggggttcca 20340  
cccctacttc gtctactcgy gctccatccc ctacctagac ggcaccttct acctcaacca 20400  
caccttcaag aaggtctcca tcaccttcca ctccctcgtc agctggcccg gcaacgaccg 20460  
cctctgagc cccaacgagt tcgaaatcaa cgcacccgcy gacggagaggy gatacaacgt 20520  
ggcccagtc aacatgacca aggactggtt cctggtccag atgctggccc actacaacat 20580  
cggctaccag ggttcttacg tgcggaggy ctacaaggac cgcagtact ccttcttccg 20640  
caacttccag cccatgagcc gccaggtcgt ggacgaggtc aactacaaggy actaccaggy 20700  
cgtcacctcy gcctaccagc acaacaactc gggcttctc ggtacctcgy cgcaccat 20760  
gcyccaggyg cagccctacc ccgccaacta cccctaccgy ctcatcgyga agagcgcct 20820  
cgcagcgtc acccagaaaa agttcctcgy cgcacgggcy atgtggcyga tccccttctc 20880  
cagaacttc atgtccatgy cgcgctcac cgcctcgyg cagaacatgc tctacgcca 20940  
ctccgcccac gcgctagaca tgaatttcca agtcgacccc atggatgagt ccaaccttct 21000  
ctatgttctc ttcgaaagtct tcgacgtcgt ccgagtgcac cagccccacc gcygctcat 21060  
cgaagcctc taactcgyga cgccttctc ggcggcaac gccaccaact aagcgcctc 21120  
tgcttcttgc aagatgacgy cgggctccgy cgcagcaggy ctcagggcca tcctccgyga 21180  
cctgggctc gggccctgct tcttgggac ctctgacaag cgttccctc gattcatggy 21240  
cccgcacaag ctggcctcgy ccatcgtgaa cacggccgyg cgcgagaccg ggggcygca 21300  
ctggtggycc ttcgcttgya acccgcgctc ccacacatgc tacctctcgy acccctcgy 21360  
gttctcgyg cgcgctca agcagatcta ccagttcgy tacgagggcy tgcgctcgy 21420  
cagcgcctc gccaccggy accgctcgt caccctgga aagtccaacc agaccgtcga 21480

---

-continued

---

gggtccgcgc tcggccgcct gcgggctctt ctgctgcatg ttctctgcacg ccttcgtgca 21540  
ctggcccgcac cgccccatgg acaagaaccc caccatgaac ttactgacgg gggtgcccaa 21600  
cgcatgctc cagtcgcccc aggtggaacc caccctgcgc cgcaaccagg aagcgctcta 21660  
ccgcttcctc aatgcccact ccgcctactt tcgctcccac cgcgcgcgca tcgagaaggc 21720  
caccgccttc gaccgcatga atcaagacat gtaaaaaacc ggtgtgtgta tgtgaatgct 21780  
ttattcataa taacagcac atgtttatgc caccttctct gaggtctga ctttatttag 21840  
aaatcgaaag gttctgcgcg gctctcggca tggcccgcgg gcagggatac gttgcggaac 21900  
tggtacttgg gcagccactt gaactcgggg atcagcagct tgggcacggg gaggtcgggg 21960  
aacgagtcgc tccacagctt gcgcgtgagt tgcagggcgc ccagcaggtc gggcgcggag 22020  
atcttgaaat cgcagttggg acccgcgttc tgcgcgcgag agttgcggta cacggggttg 22080  
cagcactgga acaccatcag gcccgggtgc ttcacgcttg ccagcacctg cgcgtcggtg 22140  
atgccctcca cgtccagatc ctcgcgcttg gccatcccga agggggcat cttgcaggtc 22200  
tgccgcccc tgctggggc gcagccgggc ttgtggttgc aatcgcagtg cagggggatc 22260  
agcatcatct gggcctgctc ggagctcatg cccgggtaca tggccttcat gaaagcctcc 22320  
agctggcggg aggcctgctg cgccttgcgc ccctcgggtg agaagacccc gcaggacttg 22380  
ctagagaact gttggtggc gcagccggcg tcgtgcacgc agcagcgcgc gtcgttggttg 22440  
gccagctgca ccacgctgcg cccccagcgg ttctgggtga tcttggcccg gttggggttc 22500  
tccttcagcg cgcgctgccc gttctcgtc gccacatcca tctcgatagt gtgctccttc 22560  
tggatcatca cggctccgtg caggcaccgc agcttgccct cggcttcggg gcagccgtgc 22620  
agccacagcg cgcagccggt gcactcccag ttcttggggg cgatctggga gtgcgagtg 22680  
acgaagccct gcaggaagcg gccatcatc gcggtcaggg tcttgttgcg ggtgaaggtc 22740  
agcgggatgc cgcggtgctc ctcgttcaaca tacaggtggc agatgcggcg gtacacctcg 22800  
ccctgctcgg gcatcagctg gaaggcggac ttcaggtcgc tctccacgcg gtaccggctc 22860  
atcagcagcg tcatcacttc catgccttc tcccaggccg aaacgatcgg caggctcagg 22920  
gggttcttca ccgccattgt catcttagtc gccgccgcgg aggtcagggg gtcgttctcg 22980  
tccagggctc caaactcgc cttgcogtcc ttctcgatga tgcgcacggg gggaaagctg 23040  
aagcccacgg ccgccagctc ctcctcggcc tgccttctgt cctcgtgctc ctggctgatg 23100  
tcttgaaaag gcacatgctt ggtcttgcgg ggttctttt tggcggcag aggcggcggc 23160  
gatgtgctgg gagagcgcga gttctcgttc accacgacta tttcttcttc ttggccgtcg 23220  
tccgagacca cgcggcggta ggcattgcctc ttctggggca gagcggagg cgacgggctc 23280  
tcgcggttcg gcggcggctt ggcagagccc cttccgcgtt cgggggtgcg ctcctggcgg 23340  
cgctgctctg actgacttcc tccgcggccg gccattgtgt tctcctaggg agcaacaaca 23400  
agcatggaga ctacgccatc gtcgccaaca tcgccatctg ccccccgcgc caccgcccac 23460  
gagaaccagc agcagaatga aagcttaacc gccccgcgc ccagcccac ctcgacgccc 23520  
gcggccccag acatgcaaga gatggaggaa tccatcgaga ttgacctggg ctaogtgacg 23580  
ccccgcggagc acgaggagga gctggcagcg cgcttttcag ccccgaaga gaaccaccaa 23640  
gagcagccag agcaggaagc agagaacgag cagaaccagg ctgggcacga gcatggcagc 23700  
tacctgagcg gggcagagga cgtgctcatc aagcatctgg cccgccaatg catcatcgtc 23760

---

-continued

---

aaggacgcgc tgctcgaccg cgcgaggtg cccctcagcg tggcggagct cagccgcgcc 23820  
tacgagcgca acctcttctc gccgcgctg ccccccaagc gccagcccaa cggcacctgt 23880  
gagcccaacc cgcgcctcaa cttctaccog gtcttcgctg tggccgaggc cctggccacc 23940  
taccacctct ttttcaagaa ccaaaggatc cccgtctcct gccgcgcaa ccgcacccgc 24000  
gccgacgccc tgctcaacct gggccccggc gccgcctac ctgatatac ctcttgtaa 24060  
gaggttccca agatcttctc gggcttggc agcgacgaga ctggggcgc gaacgctctg 24120  
caaggaagcg gagaggagca tgagcaccac agcgccttg tggagttgga aggcgacaac 24180  
gcgcgcttg cggctctcaa gcgcaggtc gagctgacc acttcgcta cccggcctc 24240  
aacctgccc ccaaggtcat gagcgcctc atggaccagg tgctcatcaa gcgcgctcg 24300  
cccctctcg aggaggagat gcaggacccc gagagttcgg acgagggcaa gccctgggtc 24360  
agcgacgagc agctggcgc ctggctggga gcgagtagca cccccagag cctggaagag 24420  
cggcgcaagc tcatgatgac cgtggtcctg gtgaccgtg agctggagtg tctgcgccc 24480  
ttctttgccc acgcgagac cctgcgcaag gtcgaggaga acctgcacta cctcttcagg 24540  
cacgggttcg tcgcccagc ctgcaagatc tccaacgtg agctgaccaa cctggctctc 24600  
tacatgggca tcctgcacga gaaccgcctg gggcaaacg tgctgcacac caccctgcgc 24660  
ggggaggccc gccgcgacta catccgcgac tgcgtctacc tgtacctctg ccacacctg 24720  
cagacgggca tggcgctgtg gcagcagtc ctggaggagc agaacctgaa agagctctgc 24780  
aagctcctgc agaagaacct caagccctg tggaccgggt tcgacgagcg taccaccgcc 24840  
tcggacctg ccgacctcat cttccccgag cgcctgcggc tgacgctgc caacgggctg 24900  
cccgacttta tgagccaaa catgttcaa aactttcgt cttctactct cgaacgctcc 24960  
gggatcctgc ccgccacctg ctccgcgctg ccctcgact tctgcccgt gacctctgc 25020  
gagtgcctc cgcgctctg gagccactg tacttgcctg gcctggcaa ctacctggc 25080  
taccactcg acgtgatcga ggacgtcagc ggcgaggtc tgctggagtg ccactgccc 25140  
tgcaacctct gcacgcgca ccgctccctg gcctgcaacc cccagctgct gagcgagacc 25200  
cagatcatcg gcacctctc gttgcaagc cccggcgagc gcgagggcaa ggggggtctg 25260  
aaactcacc cgggctgtg gacctggcc tacttgcgca agttcgtgcc cgaggactac 25320  
catcccttcg agatcaggtt ctacgagac caatcccagc cgcceaaggc cgagctgctg 25380  
gcctgcgca tcaccaggg gccatcctg gcccaattg aagccatcca gaaatcccgc 25440  
caagaatttc tgctgaaaa gggccacggc gtctacttg acccccagac cggagaggag 25500  
ctcaaccca gttccccca ggatgcccc aggaagcagc aagaagctga aagtggagct 25560  
gccgcccgc gaggatattg aggaagactg ggagagcagt caggcagagc aggaggagat 25620  
ggaagactg gcacgacac aggcagagga ggacagcctg caagacagtc tggaggagga 25680  
agacgaggtg gaggagcgag aggaagaagc agccgcccgc agaccgtcgt cctcgcgga 25740  
gaaagcaagc agcagcgata catctcccgc tccgggtcgg ggtcgcggcg gccgggcca 25800  
cagtaggtg gcagagacc ggcgctccc gaaccccacc acccagaccg gtaagaagga 25860  
gcggcaggg taacaagtct ggcgggggca caaaaacgcc atcgtctcct gcttgcaagc 25920  
ctgccccgc aacatctcct tcaccggcg ctacctgctc ttccaccgc ggggtaactt 25980  
ccccgcaac atcttgcat actacgtca cctccacagc ccctactact gtttcaaga 26040



-continued

---

```

agaggcagaa acccagcagc agcagaaaac cagcggcagc agcagctaga aaatccacag 26100
cggcggcagc tggactgagg atcgcggcga acgagccggc gcagaccggg gagctgagga 26160
accggatcct tcccaccctc tatgccatct tccagcagag tcgggggcag gacaggaac 26220
tgaaagtcaa gaaccgttct ctgcgctcgc tcacccgcag ttgtctgtat cacaagagcg 26280
aagaccaact tcagcgcact ctgcgagcgc cggaggtctt cttcaacaag tactgcgcgc 26340
tcactcttaa agagtagccc gcgccgccc acacacgaa aaaggcggga attacgtcac 26400
cacctgcgcc cttcgcccga ccatcatgag caaagagatt cccacgcctt acatgtggag 26460
ctaccagccc cagatgggcc tggccgccgg cgccgccag gactactcca cccgcatgaa 26520
ctggctcagt gccgggcccg cgatgatctc acgggtgaat gacatccgcg cccaccgaaa 26580
ccagatactc ctagaacagt cagcgatcac cgccacgccc cgccatcacc ttaatccgcg 26640
taattggccc gccgccctgg tgtaccagga aattccccag cccacgaccg tactacttcc 26700
gcgagacgcc caggccgaag tccagctgac taactcaggt gtccagctgg ccggcggcgc 26760
cgccctgtgt cgtcaccgcc ccgctcaggg tataaagcgg ctggtgatcc gaggcagagg 26820
cacacagctc aacgacgagg tggtagctc ttcgctgggt ctgcgacctg acggagtctt 26880
ccaactcgcc ggatcgggga gatcttcctt cacgcctcgt caggccgtcc tgactttgga 26940
gagttcgtcc tcgacgcccc gctcggggcg catcggcact ctccagttcg tggaggagtt 27000
cactccctcg gtctacttca accccttctc cggtccccc gccactacc cggacgagtt 27060
catcccgaac ttcgacgcca tcagcagatc ggtggacggc tacgattgaa tgtccatgg 27120
tggcgcagct gacctagctc ggcttcgaca cctggaccac tgccgccgct tccgctgctt 27180
cgctcgggat ctgcgccagt ttgcctactt tgagctgccc gaggagcacc ctccagggccc 27240
agcccacgga gtgcggatca tcgtcgaagg gggcctcgac tcccacctgc ttcggatcct 27300
cagccagcga ccgatcctgg tcgagcgcga acaaggacag acccttctta cttgtactg 27360
catctgcaac cccccggcc tgcatgaaa tctttgttgt ctgctgtgta ctgagtataa 27420
taaaagtga gatcagcgac tactcggac tcgattgtgg tgttcctgct atcaaccggt 27480
ccctgttctt caccgggaac gagaccgagc tccagctcca gtgtaagccc cacaagaagt 27540
acctcacctg gctgttccag ggctccccga tcgccgttgt caaccactgc gacaacgacg 27600
gagtcctgct gagcggccct gccaacctta cttttccac ccgcagaagc aagctccagc 27660
tcttcaacc ctctcctccc gggacctatc agtgcgctc aggacctgc catcacacct 27720
tccacctgat ccgaatacc acagcggcg tccccgtac taacaaccaa actaccacc 27780
aacgccaccg tcgagacctt tcctctgaat ctaataccac taccgagggt gagctccgag 27840
gtcgaccaac ctctgggatt tactacggcc cctgggaggt ggtggggtta atagcgtag 27900
gcctagttgc ggggtggctt ttggttctct gctacctata cctccctgc tgttcgtact 27960
tagtggtgct gtgttgctgg tttaagaaat ggggaagatc accctagtga gctgcggtgc 28020
gctggtggcg gtgttgcttt cgattgtggg actggggcgc gcggctgtag tgaaggagaa 28080
ggcgatccc tgcttgcat tcaatccaa caaatgccag ctgagttttc agcccgatgg 28140
caatcggtgc gcggtactga tcaagtgcg atgggaatgc gagaactga gaatcgagta 28200
caataacaag actcggaaac atactctcgc gtccgtgtgg cagcccgggg accccgagtg 28260
gtacaccgct tctgtccccg gtgctgacgg ctccccgcgc accgtgaata atactttcat 28320

```

---

-continued

---

ttttgcgcac atgtgcaaca cggatcatgt gatgagcaag cagtacgata tgtggcccc 28380  
cacgaaggag aacatcgtgg tcttctccat cgcttacagc ctgtgcacgg cgtaaatcac 28440  
cgctatcgtg tgcttgagca ttcacatgct catcgctatt cgtccagaa ataatgccga 28500  
gaaagagaaa cagccataac acgttttttc acacaccttg tttttacaga caatgcgtct 28560  
gttaaatttt ttaaacattg tgctcagtat tgcttatgcc tctggttatg caaacatata 28620  
gaaaaccctt tatgtaggat ctgatggtac actagagggt acccaatcac aagccaaggt 28680  
tgcatggtat ttttatagaa ccaacactga tccagttaaa ctttgtaagg gtgaattgcc 28740  
gcgtacacat aaaactccac ttacatttag ttgcagcaat aataatctta cacttttttc 28800  
aattacaaaa caatatactg gtacttatta cagtacaaac tttcatacag gacaagataa 28860  
atattatact gttaaaggtag aaaatcctac cactcctaga actaccacca ccaccactac 28920  
tgcaaagccc actgtgaaaa ctacaactag gaccaccaca actacagaaa ccaccaccag 28980  
cacaacactt gctgcaacta cacacacaca cactaagcta accttacaga ccactaatga 29040  
tttgatcgcc ctgctgcaaa agggggataa cagcaccact tccaatgagg agatacccaa 29100  
atccatgatt ggcattattg ttgctgtagt ggtgtgcatg ttgatcatcg ccttgtgcat 29160  
ggtgtactat gccttctgct acagaaagca cagactgaac gacaagctgg aacacttact 29220  
aagtgttgaa ttttaatttt ttagaacat gaagatccta ggccttttta gttttctat 29280  
cattacctct gctctttgtg aatcagtgga tagagatgtt actattacca ctggttctaa 29340  
ttatacactg aaaggccac cctcaggtat gcttctgtgg tattgctatt ttggaactga 29400  
cactgatcaa actgaattat gcaattttca aaaaggcaaa acctcaaact ctaaaatctc 29460  
taattatcaa tgcaatgga ctgatctgat actactcaat gtcacgaaag catatggtgg 29520  
cagttattat tgccctggac aaaacactga agaatgatt ttttcaaag tggaagtgg 29580  
tgatccact acaccacca ccaccacaac tattcatacc acacacacag aacaaacacc 29640  
agaggcaaca gaagcagagt tggccttcca ggttcacgga gattcctttg ctgtcaatac 29700  
ccctacacc gatcagcggg gtcggggccc gctagtcagc ggcatgtcg gtgtgctttc 29760  
gggattagca gtcataatca tctgcatgtt catttttctg tgctgtata gaaggcttta 29820  
ccgacaaaa tcagaccac tgctgaacct ctatgtttaa tttttccag agcoatgaag 29880  
gcagttagcg ctctagtttt ttgttctttg attggcattg tttttaatag taaaattacc 29940  
agagttagct ttattaaaca tgtaaatgta actgaaggag ataacatcac actagcaggt 30000  
gtagaagggt ctcaaacac cacctggaca aaataccatc taggatggag agatatttgc 30060  
acctggaatg taacttatta ttgcatagga gttaatctta ccattgttaa cgtaaccaa 30120  
tctcagaatg ggttaattaa aggacagagt gttagtgtga ccagtgatg gtactatacc 30180  
cagcatagtt ttaactacaa cattaactgc ataccactgc ctacgcctag cccactagc 30240  
actaccacac agacaaccac atacagtaca tcaaatcagc ctaccaccac tacagcagca 30300  
gaggttgcca gctcgtctgg ggtccgagtg gcatttttga tgttgcccc atctagcagt 30360  
cccactgcta gtaccaatga gcagactact gaatttttgt ccactgtoga gagccacacc 30420  
acagctacct ccagtgctt ctctagcacc gccaatctct cctcgtttc ctctacacca 30480  
atcagcccc ctactactcc tagccccgt cctcttcca ctcccctga gcaaacagac 30540  
ggcggcatgc aatggcagat caccctgctc attgtgatcg ggttgatcat cctggcctg 30600

---

-continued

---

ttgctctact acatcttctg cgcgccatt cccaacgcgc accgcaagcc ggcctacaag 30660  
cccatacgta tcgggcagcc ggagccgctt caggtggaag ggggtctaag gaatcttctc 30720  
ttctctttta cagtatggtg attgaaactat gattcctaga caattcttga tcaactattct 30780  
tatctgcctc ctccaagtct gtgccacct cgctctggtg gccaacgcca gtccagactg 30840  
tattgggccc ttgcctcct acgtgctctt tgccttcgtc acctgcatct gctgctgtag 30900  
catagtctgc ctgcttatca ccttcttcca gttcattgac tggatctttg tgcgcatcgc 30960  
ctacctgcgc caccaccccc agtaccgca ccagcgagtg gcgagctgc tcaggctcct 31020  
ctgataagca tgcggctct gctacttctc gcgcttctgc tgttagtget cccccgtccc 31080  
gtcgaccccc ggtccccac tcagtcccc gaggaggttc gcaaatgcaa attccaagaa 31140  
ccctgaaat tctcaaatg ctaccgcaa aatcagaca tgcattcccag ctggatcatg 31200  
atcattggga tcgtgaacat tctggctgc accctcatct cctttgtgat ttaccctgc 31260  
tttgactttg gttgaaactc gccagagcg ctctatctcc gcctgaacc tgacacacca 31320  
ccacagcagc aacctcaggc acacgcacta ccaccaccac agcctaggcc acaatacatg 31380  
cccatattag actatgagcg cgagccacag cgaccatgc tccccgtat tagttacttc 31440  
aatctaaccg gcggagatga ctgaccact ggccaataac aacgtcaacg accttctcct 31500  
ggacatggac ggccgcgct cggagcagcg actcgccaa cttcgattc gtcagcagca 31560  
ggagagagcc gtcaaggagc tgcaggacgg catagccatc caccagtga agagaggcat 31620  
cttctgcctg gtgaaacag ccaagatctc ctacgagtc accagaccg accatcgct 31680  
ctctacgag ctctgcagc agcgcagaa gttcacctgc ctggtcggag tcaacccat 31740  
cgtcatcacc cagcagtcgg gcgatacaca ggggtgcatc cactgctcct gcgactccc 31800  
cgactgcgtc cacactctga tcaagacct ctgcggcctc cgcgacctc tccccatgaa 31860  
ctaatacccc cttatccag tgaataaag atcatattga tgatgattta aataaaaaa 31920  
ataatcattt gatttgaaat aaagataca tcatattgat gatttgagtt taacaaaaat 31980  
aaagaatcac ttacttgaaa tctgatacca ggtctctgtc catgtttctt gccaacacca 32040  
ctcactccc ctcttcccag ctctggtact gcaggccccg gcgggctgca aacttctctc 32100  
acacgctgaa ggggatgta aattcctcct gtccctcaat cttcatttta tcttctatca 32160  
gatgtccaaa aagcgcgctc ggggtgatga tgacttcgac cccgtctacc cctacgatgc 32220  
agacaacgca ccgaccgtgc cttcatcaa ccccccttc gtctctcag atggattcca 32280  
agagaagccc ctgggggtgt tctccctcg actggctgac cccgtcaca ccaagaacgg 32340  
ggaaatcacc ctcaagctgg gagaggggtt ggacctcgac tcgtcgggaa aactcatctc 32400  
caacacggcc accaaggccg ccgccctct cagtatttca aacaacacca tttccctta 32460  
aactgctgcc ctttctaca acaacaatgg aactttaagc ctcaatgtct ccacaccatt 32520  
agcagtattt cccacattta acactttagg cataagtctt ggaacggtc ttcagacttc 32580  
aaataagttg ttgactgtac aactaactca tctcttaca ttcagctcaa atagcatcac 32640  
agtaaaaaca gacaaagggc tataatattaa ctccagtggg aacagaggac ttgaggctaa 32700  
tataagccta aaaagaggac tagtttttga cggtaatgct attgcaacat atattgaaa 32760  
tggcttagac tatggatctt atgatagtga tggaaaaaca agaccgtaa ttacaaaaat 32820  
tggagcagga ttaaattttg atgctaacaa agcaatagct gtcaactag gcacaggttt 32880

---

-continued

---

aagttttgac tccgctggtg ccttgacagc tggaaacaaa caggatgaca agctaacact 32940  
ttggactacc cctgacccaa gccctaattg tcaattactt tcagacagag atgccaaatt 33000  
tactctctgt cttacaaaat gcggtagtca aatactaggc actgtggcag tggcggctgt 33060  
tactgtagga tcagcactaa atccaattaa tgacacagtc aaaagcgcca tagttttcct 33120  
tagatattgat tccgatggtg tactcatgtc aaactcatca atggtagggtg attactggaa 33180  
ctttagggag ggacagacca ctcaaagtgt agcctataca aatgctgtgg gattcatgcc 33240  
aaatataggt gcatatccaa aaacccaaag taaaacacct aaaaatagca tagtcagtca 33300  
ggtatattta actggagaaa ctactatgcc aatgacacta accataactt tcaatggcac 33360  
tgatgaaaaa gacacaaccc cagttagcac ctactctatg acttttacct ggacgtggac 33420  
tggagactat aaggcaaaaa atattacctt tgctaccaac tcattctctt tttcctacat 33480  
cgcccaggaa taatcccacc cagcaagcca accccttttc ccaccacctt tgtctatatg 33540  
gaaactctga aacagaaaaa taaagttcaa gtgttttatt gaatcaacag ttttacagga 33600  
ctcgagcagt ttttttctct ccaccctccc aggacatgga atacaccacc ctctcccccc 33660  
gcacagcctt gaacatctga atgccattgg tgatggacat gcttttggtc tccacgttcc 33720  
acacagtttc agagcgagcc agtctcgat cggtcaggga gatgaaaccc tccgggcact 33780  
cccgcactct cacctcacag ctcaacagct gaggattgtc ctcggtggtc gggatcacgg 33840  
ttatctggaa gaagcagaag agcggcggtg ggaatcatag tccgcgaacg ggatcgccg 33900  
gtggtgtcgc atcaggcccc gcagcagtcg ctgccgcgcg cgctccgtca agctgctgct 33960  
caggggggtc gggtcagggt actccctcag catgatgcc acggccctca gcatcagtcg 34020  
tctggtgctg cgggcgcagc agcgcagtcg aatctcgctc aggtcactgc agtacgtgca 34080  
acacaggacc accaggttgt tcaacagtcc atagttcaac acgctccagc cgaactcat 34140  
cgcgggaag atgctaccca cgtggcctgc gtaccagatc ctacaggtaaa tcaagtggcg 34200  
ctccctccag aagacgctgc ccatgtacat gatctccttg ggcagtggtc ggttaccac 34260  
ctcccgtac cacatcacc tctggtgaa catgcagccc cggatgatcc tgcggaacca 34320  
cagggccagc accgccccgc ccgcatgca gcgaagagac cccggatccc ggcaatgaca 34380  
atggaggacc caccgctcgt acccgtggat catctgggag ctgaacaagt ctatgttggc 34440  
acagcacagg catatgctca tgcactctc cagcactctc agctcctcgg gggtaaaaac 34500  
catatcccag ggcacgggga actcttgag gacagcgaac cccgcagaac agggcaatcc 34560  
tcgcacataa cttacattgt gcatggacag ggtatcgcaa tcaggcagca ccgggtgatc 34620  
ctccaccaga gaagcgcggg tctcggctc ctacacagct ggtaaggggg ccggcccata 34680  
cgggtgatgg cgggacgcgg ctgatcgtgt tctcgaccgt gtcagatgc agttgcttcc 34740  
ggacattttc gtacttctgt tagcagaacc tggtcggggc gctgcacacc gatcgccggc 34800  
ggcgtctcgc gcgcttgaa cgctcgggtg taaagttgta aaacagccac tctctcagac 34860  
cgtgcagcag atctagggcc tcaggagtga tgaagatccc atcatgcctg atagctctga 34920  
tcacatcgac caccgtgaa tgggcccagg ccagccagat gatgcaattt tgttgggttt 34980  
cggtgacggc gggggaggga agaacaggaa gaaccatgat taacttttaa tccaaacggt 35040  
ctcgagcacc ttcaaaatga aggtcacgga gatggcacct ctgcccccg ctgtgttgg 35100  
ggaaaataac agccaggtca aaggtgatac ggttctcgag atgttccacg gtggcttcca 35160

-continued

---

```

gcaaagcctc cacgcgcaca tccagaaaca agacaatagc gaaagcggga gggttctcta 35220
atctctcaac catcatgta cactcctgca ccatccccag ataattttca tttttccage 35280
cttgaatgat tcgaactagt tcctgaghta aatccaagcc agccatgata aaaagctcgc 35340
gcagagcacc ctcccaccgc attcttaagc acaccctcat aattccaaga tattctgctc 35400
ctggttcacc tgcagcagat tgacaagcgg aatatcaaaa tctctgccgc gatccctgag 35460
ctctccctc agcaataact gtaagtaact tttcatatcg tctccgaaat ttttagccat 35520
aggaccccc ggaataagag aagggcaagc cacattacag ataaccgaa gtcccccca 35580
gtgagcattg ccaaagttaa gattgaaata agcatgctgg ctagaccggg tgatatcttc 35640
cagataactg gacagaaaat cgggtaagca attttaaga aaatcaaca aagaaaaatc 35700
ttccagggtc acgtttaggg cctcgggaac aacgatggag taagtgaag ggggtcggtc 35760
cagcatgggt agttagctga tctgtaaaaa aacaaaaaat aaaacattaa accatgctag 35820
cctggcgaac aggtgggtaa atcgttctct ccagcaccag gcaggccacg gggctcgg 35880
cgcgaccctc gtaaaaattg tcgctatgat tgaaaacat cacagagaga cgttccgggt 35940
ggcggcgggt aatgattcga gaagaagcat acacccccgg aacattggag tccgtgagtg 36000
aaaaaaagcg gccgaggaag caatgaggca ctacaacgct cactctcaag tccagcaaag 36060
cgatgccatg cggatgaagc aaaaaatctt caggtgcgta aaaaatgtaa ttactcccct 36120
cctgcacagc cagcgaagct cccgatccct ccagatacac atacaaagcc tcagcgtcca 36180
tagcttaccg agcggcagca gcagcggcac acaacaggcg caagagtcag agaaaagact 36240
gagctctaac ctgtccggcc gctctctgct caatatatag cccagatct aactgacgt 36300
aaaggccaaa gtctaaaaat acccgccaaa taatcacaca cgcccagcac acgcccagaa 36360
accggtgaca cactcagaaa aatacgcgca cttcctcaa cgccaaact gccgtcattt 36420
ccgggttccc acgctacgtc atcaaacac gactttcaa ttccgtcgac cgttaaaaac 36480
atcacccgcc ccgcccctaa cggtcgccgc tcccgcagcc aatcaccttc ctccctccc 36540
aaattcaaac agctcatttg catattaacg cgcacccaaa gtttgaggta tattattgat 36600
gatg 36604

```

&lt;210&gt; SEQ ID NO 3

&lt;211&gt; LENGTH: 36535

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: chimpanzee adenovirus serotype Pan7

&lt;400&gt; SEQUENCE: 3

```

catcatcaat aatatacctc aaacttttgg tgcgcgtaaa tatgcaaatg agctgtttga 60
atctggggag ggaggaaggt gattggccga gagacgggag accgttaggg gcggggcggg 120
tgactttttt aatacgtggc cgtgaggcgg agccggtttg caagtctctg tgggaaaagt 180
gacgtcaaac gaggtgtggt ttgaacacgg aaatactcaa tttcccgcgc ctctctgaca 240
ggaaatgagg tgtttctggg cggatgcaag tgaaaacggg ccattttcgc gcgaaaactg 300
aatgaggaag tgaaaatctg agtaatttcg cgtttatggc agggaggagt atttgccgag 360
ggccagtag actttgaccg attacgtggg ggtttcgatt accgtatctt tcacctaaat 420
ttccgcgtac ggtgtcaaa gtcgggtgtt ttacgtaggc gtcagctgat cgcagggta 480
tttaaacctg cgctctctag tcaagaggcc actcttgagt gccagcagat agagttttct 540

```

-continued

---

|  |      |
|--|------|
| cctccgcgcc gcgagtcaga tctacacttt gaaagatgag gcacctgaga gacctgccc   | 600  |
| gtaatgtttt cctggtactt gggaacgaga ttctggaatt ggtggtggac gccatgatgg  | 660  |
| gtggcgaccc tcctgagccc cctaccccatt ttgagggccc ttcgctgtac gatttgtatg | 720  |
| atctggaggt ggatgtgccc gagaacgacc ccaacgagga ggcggtgaat gatttgttta  | 780  |
| gcgagccgc gctgctggct gccgagcagg ctaatacggg ctctggctca gacagcgatt   | 840  |
| cctctctcca taccocgaga cccggcagag gtgagaaaa gatccccgag cttaaagggg   | 900  |
| aagagctcga cctgcctgct tatgaggaat gcttgcctcc gagcgatgat gaggaggacg  | 960  |
| aggaggcgtat tcgagctgca tcgaaccagg gagtgaaagc tgcgggcgaa agctttagcc | 1020 |
| tggactgtcc tactctgccc ggacacggct gtaagtcttg tgaatttcat cgcatgaata  | 1080 |
| ctggagataa gaatgtgatg tgtgccctgt gctatatgag agcttacaac catttgtttt  | 1140 |
| acagtaagtg tgattaactt tagttgggaa ggcagagggg gactgggtgc tgactggttt  | 1200 |
| atztatgtat atgttttttt atgtgtaggt cccgtctctg acgtagatga gacccccact  | 1260 |
| tcagagtgca tttcatcacc cccagaaatt ggcgaggaac cgcccgaaga tattattcat  | 1320 |
| agaccagttg cagtgtgagt caccgggccc agagcagctg tggagagttt ggatgacttg  | 1380 |
| ctacagggty gggatgaacc tttggacttg tgtaccggga aacgcccag gactaagtg    | 1440 |
| ccacacatgt gtgtttactt aagggtgatg cagtatttat aggggtgtgga gtgcaataaa | 1500 |
| atccgtgttg actttaagtg cgtggtttat gactcagggg tggggactgt gggatatata  | 1560 |
| gcaggtgcag acctgtgtgg tcagttcaga gcaggactca tggagatctg gacggtcttg  | 1620 |
| gaagactttc acccagactag acagctgcta gagaactcat cggagggggg ctcttacctg | 1680 |
| tggagattct gcttcggtgg gcctctagct aagctagtct atagggccaa acaggattat  | 1740 |
| aaggatcaat ttgagatat tttgagagag tgtccctggta tttttgactc tctcaacttg  | 1800 |
| ggccatcagt ctcaacttaa ccagagtatt ctgagagccc ttgacttttc tactcctggc  | 1860 |
| agaactaccg ccgcggtagc cttttttgcc tttatccttg acaaatggag tcaagaacc   | 1920 |
| catttcagca gggattaccg tctggactgc ttagcagtag ctttgtggag aacatggagg  | 1980 |
| tgccagcgcc tgaatgcaat ctccggctac ttgccagtag agccggtaga cacgctgagg  | 2040 |
| atcttgatc tccagtcacc ccaggaacac caacgccgcc agcagccgca gcaggagcag   | 2100 |
| cagcaagagg aggaggagga tcgagaagag aaccgagag ccggtctgga ccctccggtg   | 2160 |
| gcggaggagg aggagtagct gacttgtttc ccgagctgcg ccgggtgctg actaggctct  | 2220 |
| ccagtgagc ggagagggg attaagcggg agaggcatga ggagactagc cacagaactg    | 2280 |
| aactgactgt cagtctgatg agccgcagc gcccagaatc ggtgtggtg catgaggttc    | 2340 |
| agtgcaggg gatagatgag gtctcgggta tgcagtagaa atattccctg gaacaagtca   | 2400 |
| agacttgttg gttggagcct gaggatgatt gggaggtagc catcaggaat tatgccaagc  | 2460 |
| tggctctgaa gccagacaag aagtacaaga ttaccaaact gattaatattc agaaattcct | 2520 |
| gtacatttc agggaatggg gccgaggtgg agatcagtag ccaggagagg gtggccttca   | 2580 |
| gatgttgat gatgaatatg taccggggg tgggtggcat ggaggagtc acctttatga     | 2640 |
| acgcgaggtt caggggtgat gggataatg ggggtgctt tatggccaac accaagctga    | 2700 |
| cagtgcacgg atgctccttc tttgggttca ataacatgtg catcgaggcc tggggcagtg  | 2760 |
| tttcagtgag gggatgcagc ttttcagcca actggatggg ggtcgtgggc agaaccaaga  | 2820 |

-continued

---

|  |      |
|--|------|
| gcaaggtgtc agtgaagaaa tgccctgttc agaggtgcc a cctgggggtg atgagcgagg | 2880 |
| gcgaagccaa agtcaaacac tgcgcctcta ctgagacggg ctgctttgtg ctgatcaagg  | 2940 |
| gcaatgccca agtcaagcat aacatgatct gtggggcctc ggatgagcgc ggctaccaga  | 3000 |
| tgctgacctg cgccgggtgg aacagccata tgctggccac cgtgcatgtg acctcgcacc  | 3060 |
| cccgaagac atggcccag ttcgagcaca acgtcatgac ccgatgcaat gtgcacctgg    | 3120 |
| ggccccgcg aggcattgtc atgccctacc agtgcaacat gcaatttgtg aagggtctgc   | 3180 |
| tgagaccga tgccatgtcc agagtgagcc tgacgggggt gtttgacatg aatgtggagc   | 3240 |
| tgtgaaaaat tctgagatat gatgaatcca agaccagtg ccgggcctgc gaatgcggag   | 3300 |
| gcaagcacgc caggcttcag ccctgtgtgt tggagggtgac ggaggacctg cgaccgatc  | 3360 |
| atgtgtgtt gtccctgcaac gggacggagt tcggctccag cggggaagaa tctgactaga  | 3420 |
| gtgagtgtg tttgggggag gtggagggtc tgtatgaggg gcagaatgac taaaatctgt   | 3480 |
| gtttttctgt gtgttgacg agcatgagcg gaagcgcctc ctttgaggga ggggtattca   | 3540 |
| gcccttatct gacggggcgt ctcccctcct gggcgggagt gcgtcagaat gtgatgggat  | 3600 |
| ccacggtgga cgccggccc gtgcagccc cgaactctc aacctgacc tacgcgacc       | 3660 |
| tgagctcctc gtccgtggac gcagctgcc ccgcagctgc tgcttccgcc gccagcgcg    | 3720 |
| tgcgcggaat ggccctgggc gccggctact acagctctct ggtggccaac tcgactcca   | 3780 |
| ccaataatcc cgccagcctg aacgaggaga agctgctgct gctgatggcc cagctcgagg  | 3840 |
| ccctgaccga gcgcctgggc gagctgacct agcaggtggc tcagctgacg gcggagacgc  | 3900 |
| ggcccgcggt tgccacgggt aaaacaaat aaaaaatgaa tcaataaata aacgggagcg   | 3960 |
| gtgtttgatt ttaacacaga gtcttgaatc tttatgtgat ttttcgcgcg cggtagggcc  | 4020 |
| tggaaccacc gtctcgatca ttgagcacc ggtggatgtt ttccaggacc cggtagaggt   | 4080 |
| ggccttggat gttgaggtac atgggcatga gcccgctccg ggggtggagg tagctccatt  | 4140 |
| gcagggcctc gtgctcgggg gtggtgtgtt aatcaccga gtcatagcag gggcgcaggg   | 4200 |
| cgtggtgctg cacgatgtcc ttgaggagga gactgatggc cacgggcagc cccttgggtg  | 4260 |
| aggtgttgac gaacctgttg agctgggagg gatgcatgcg gggggagatg agatgcatct  | 4320 |
| tgccctggat cttgagattg gcgatgttcc cgcccagatc ccgcccgggg ttcatgttgt  | 4380 |
| gcaggaccac cagcacggtg tatccggtgc acttggggaa tttgtcatgc aacttggaa   | 4440 |
| ggaaggcgtg aaagaatttg gagacgccct tgtgaccgcc caggttttcc atgcaactcat | 4500 |
| ccatgatgat ggcgatgggc ccgtggggcg cggcctgggc aaagacgttt cgggggtcgg  | 4560 |
| acacatcgta gttgtgttcc tgggtgagct cgtcataggc cattttaatg aatttggggc  | 4620 |
| ggaggggtgc cgactggggg acgaagggtc cctcgatccc gggggcgtag ttgccctcgc  | 4680 |
| agatctgcat ctcccaggcc ttgagctcgg agggggggat catgtccacc tgcggggcga  | 4740 |
| tgaaaaaac ggtttccggg gcgggggaga tgagctgggc cgaagcagg ttccggagca    | 4800 |
| gctgggactt gccgcagccg gtggggccgt agatgacccc gatgaccggc tgcagggtgt  | 4860 |
| agttgagggg gagacagctg ccgtcctcgc ggaggagggg gccaccctcg ttcatcatct  | 4920 |
| cgcgacatg catgttctcg cgcacgagtt ccgccaggag gcgctcgcgc ccagcagaga   | 4980 |
| ggagctcttg cagcgaggcg aagtttttca gcggcttgag yccgtcggcc atgggcattt  | 5040 |
| tggaaggggt ctgttgcaag agttccagac ggtcccagag ctcggtgatg tgctctaggg  | 5100 |

---

-continued

---

catctcgatc cagcagacct cctcgtttcg cgggttgggg cgactgcggg agtagggcac 5160  
caggcgatgg gcgtccacgg aggccagggt ccggtccttc cagggtcgca gggcccgct 5220  
cagcgtggtc tccgtcacgg tgaaggggtg cgcgccgggc tgggcgcttg cgagggtgcg 5280  
cttcaggctc atccggctgg tcgagaaccg ctcccggctg gcgccctgcg cgtcggccag 5340  
gtagcaattg agcatgagtt cgtagttgag cgcctcggcc gcgtggccct tggcgcggag 5400  
cttacctttg gaagtgtgtc cgcagacggg acagaggagg gacttgaggg cgtagagctt 5460  
gggggcgagg aagacggact cggggcgcta ggcgtccgcg ccgacagctg cgcagacggg 5520  
ctcgcactcc acgagccagg tgaggtcggg ccggttgggg tcaaaaacga ggtttcctcc 5580  
gtgctttttg atgctttct tacctctggt ctccatgagc tcgtgtcccc gctgggtgac 5640  
aaagaggctg tccgtgtccc cgtagaccga ctttatgggc cggtcctcga gcggggtgcc 5700  
gcggtcctcg tcgtagagga accccgccca ctccgagacg aagccccggg tccaggccag 5760  
cacgaaggag gccacgtggg aggggtagcg gtcgttgtcc accagcgggt ccacctctc 5820  
cagggtatgc aagcacatgt ccccctcgtc cacatccagg aaggtgattg gcttgaagt 5880  
gtaggccacg tgaccggggg tcccggccgg gggggtataa aagggggcgg gccctgctc 5940  
gtcctcactg tcttccggat cgctgtccag gagcgcagc tgttgggta ggtattccct 6000  
ctcgaaggct gccataacct cggcactcag gttgtcagtt tctagaaacg aggaggattt 6060  
gatattgacg gtgccgttgg agacgccttt catgagcccc tcgtccatct ggtcagaaaa 6120  
gacgatcttt ttgtgtcga gcttgggtgc gaaggagccg tagaggcgt tggagaggag 6180  
cttggcgatg gagcgcattg tctggttctt ttccttctg gcgcgctcct tggcggcgat 6240  
gttgagctgc acgtactcgc gcgccacgca cttccattcg gggaaagcgg tggtagctc 6300  
gtcgggcacg attctgacct gccagcccg gttgtgcagg gtgatgagg ccacgctggt 6360  
ggccacctcg ccgcgcagg gctcgttgg ccagcagagg cgcgccctc tgcgcgagca 6420  
gaaggggggc agcgggtcca gcatgagctc gtcggggggg tcggcgtcca cggtaagat 6480  
gccgggcaga agctcggggg cgaagtagct gatgcagggt tccagatcgt ccagcgcgc 6540  
ttgacagctg cgcacggcca gcgcgcgctc gttagggctg aggggcgtgc cccaggcat 6600  
ggggtcgtg agcgcggagg cgtacatgcc gcagatgctg tagacgtaga ggggtcctc 6660  
gaggacgccg atgtaggtgg ggtagcagcg cccccgcgg atgctggcgc gcacgtagtc 6720  
gtacagctcg tcgagggcg cgaggagccc cgtgccagg ttggagcgtt gcggctttc 6780  
ggcgcgtag acgatctgac ggaagatggc gtgggagttg gaggagatgg tggcctctg 6840  
gaagatggtg aagtggcgt ggggcagcc gaccgagtcc ctgatgaagt gggcgtagga 6900  
gtcctgcagc ttggcgacga gctcggcggg gacgaggacg tccagggcgc agtagtcgag 6960  
ggtctcttgg atgatgctg acttgagctg gcccttctg tccacagct cgcggttgg 7020  
aaggaactct tcgcgctcct tccagtactc ttcgaggggg aaccgctcct gatcggcagc 7080  
gtaagagccc accatgtaga actggttgc ggcctttag gcgcagcagc ccttctccac 7140  
ggggaggggc taagcttgtg cggccttgcg cagggaggtg tgggtgaggg cgaaggtgtc 7200  
gcgcaccatg acctgagga actggtgctt gaagtcgagg tcgtcgcagc cgcctgctc 7260  
ccagagctgg aagtccgtgc gcttcttcta ggcggggtt ggcaaacga aagtaacatc 7320  
gttgaagagg atcttgcggc cgcggggcat gaagttgcga gtgatgcgga aaggctgggg 7380



---

-continued

---

cacctcggcc cggttgttga tgacctgggc ggcgaggacg atctcgtcga agccgttgat 7440  
gttgtgcccc acgatgtaga gttccacgaa tcgcgggcgg cccttaacgt ggggcagctt 7500  
cttgagctcg tcgtaggtga gctcggcggg gtcgctgagc ccgtgctgct cgagggccca 7560  
gtcggcgacg tgggggttgg cgctgaggaa ggaagtccag agatccacgg ccagggcgggt 7620  
ctgcaagcgg tcccggctact gacggaaactg ctggcccacg gccatTTTTT cgggggtgac 7680  
gcagtagaag gtgcggggggt cgccctgcca gcggtccac ttgagctgga ggcgagggtc 7740  
gtgggcgagc tcgacgagcg gcgggtcccc ggagagtttc atgaccagca tgaaggggac 7800  
gagctgcttg ccgaaggacc ccacccaggt gtaggtttcc acatcgtagg tgaggaagag 7860  
cctttcggtg cgaggatgcg agccgatggg gaagaactgg atctcctgcc accagttgga 7920  
ggaatggctg ttgatgtgat ggaagtagaa atgccgacgg cgcgccgagc actcgtgctt 7980  
gtgtttatac aagcgtccgc agtgctgcga acgctgcacg ggatgcacgt gctgcacgag 8040  
ctgtacctgg gttcctttga cgaggaattt cagtgggcag tggagcctg gcggctgcat 8100  
ctggtgctgt actacgtcct gcccatcgcc gtggccatcg tctgcctcga tgggtgcat 8160  
gctgacgagc ccgcccggga ggcaggtcca gacttcggct cggacgggtc ggagagcag 8220  
gacgagggcg cgcagcccg agctgtccag ggtcctgaga cgctgcggag tcaggtcagt 8280  
gggcagcggc ggccgcccgt tgaactgacg gagcttttcc agggcgcgcg ggaggtccag 8340  
atggtacttg atctccacgg cgcctgttgg ggcgacgtcc acggttga gggccccgtg 8400  
ccccggggc gccaccaccg tgcccgtttt cttcttggg gctgcttcca tgccggtcag 8460  
aagcggcggc gaggacgcgc gccggcggc aggggcggct cgggaccggg aggcaggggc 8520  
ggcaggggca cgtcggcgcc gcgcccggc aggttctggt actgcgcccg gagaagactg 8580  
gcgtgagcga cgacgcgacg gttgacgtcc tggatctgac gcctctgggt gaaggccacg 8640  
ggaccctgta gtttgaacct gaaagagagt tcgacagaat caatctcgggt atcgttgacg 8700  
gcggcctgcc gcaggatctc ttgcaogtcg cccgagttgt cctggtaggc gatctcggtc 8760  
atgaactgct cgatctcctc ctcctgaagg tctccgccc cgcgcccgtc gacggtggcc 8820  
gcgaggtcgt tggagatgcg gcccatgacg tgcgagaagg cgttcctgccc gccctcgttc 8880  
cagacggcgc tgtagaccac ggctcogtcg gggtcgccc cgcgcatgac cacctgggcg 8940  
aggttgagct cgacgtggcg cgtgaagacc gcgtagttgc agaggcgtg gtagaggtag 9000  
ttgagcgtgg tggcagatg ctcggtgacg aagaagtaca tgatccagcg gcggagcggc 9060  
atctcgtcga cgtcggccag ggcttccaag cgctccatgg cctcgtagaa gtcaccggcg 9120  
aagttgaaaa actgggagtt gcgcccggc acggtcaact cctcctccag aagacggatg 9180  
agctcagcga tgggtggcgc cacctcgcgc tcgaaggccc cggggggctc ctcttcttc 9240  
atctcttct cctccactaa catctcttct acttctctct caggagcggc cggcggggga 9300  
ggggccctgc gtcgcccggc gcgcaaggc agacggtcga tgaagcgtc gatggtctcc 9360  
ccgcccggc gacgcatggt ctcggtgacg gcgcccctg cctcggggg ccgcagcgtg 9420  
aagacggcgc cgcgcatctc caggtggcgc cggggggggt ctcggttggg cagggagagg 9480  
gcgctgacga tgcatttat caattggccc gtagggactc cgcgcaagga cctgagcgtc 9540  
tcgagatcca cgggatccga aaaccgctga acgaaggctt cgagccagtc gcagtcgcaa 9600  
ggtaggctga gcccggttcc ttgttcttcg gggatttcg gaggcgggcg ggcgatgctg 9660

---

-continued

---

ctggtgatga agttgaagta ggcggtcctg agacggcggga tggtagcgag gagcaccagg 9720  
tccttggggc cggcttgctg gatgcgcaga cggtcggcca tgccccaggc gtggtcctga 9780  
cacctggcga ggtccttgta gtagtcctgc atgagccgct ccacgggcac ctctcctcgc 9840  
cccgcgcggc cgtgcatcgc cgtgagcccg aaccgcgct ggggctggac gagcgcagg 9900  
tcggcgacga cgcgctcggc gaggatggcc tgctgtatct gggtaggggt ggtctggaag 9960  
tcgtcgaagt cgacgaagcg gtggtaggct ccggtgttga tggatatagga gcagttggcc 10020  
atgacggacc agttgacggt ctggtggcgc ggtcgacga gctcgtggta cttgaggcgc 10080  
gagtaggggc gcgtgtcgaa gatgtagtcg ttgcagggtc gcacgaggta ctggtatccg 10140  
acgaggaagt ggcgcggcgc ctggcggtag agcggccatc gctcggggc gggggcgcgc 10200  
ggcgcgaggc cctcagcat gaggcggtag tagccgtaga tgtacctgga catccagggtg 10260  
atgccggcgc cgggtgtgga ggcgcgcggc aactcgcgga cgcggttcca gatggtgcgc 10320  
agcggcagga agtagttcat ggtggccgcg gtctggcccg tgaggcgcgc gcagtcgtgg 10380  
atgctctaga catacgggca aaaacgaaag cggtcagcgg ctcgactccg tggcctggag 10440  
gctaagcgaa cgggttgggc tgcgcgtgta ccccggtcgc aatctcgaat caggctggag 10500  
ccgcagctaa cgtggtactg gcaactccctc ctcgacccaa gcctgctaac gaaacctcca 10560  
ggatacggag gcgggtcgtt ttttggcctt ggtcgtcgtt catgaaaaac tagtaagcgc 10620  
ggaaagcgac cccccgcgat ggctcgtcgc cgtagctcgg agaaagaatc gccagggttg 10680  
cgttgccggtg tgccccggtt cgagcctcag cgctcggcgc cggccggtt ccgcggctaa 10740  
cgtggcgctg gctgccccgt cgtttccaag accccttagc cagccgactt ctccagttac 10800  
ggagcgagcc cctctttttt ttgtgttttt gccagatgca tcccgtactg cggcagatgc 10860  
gccccacc cccacctcaa ccgcccctac cgcgcgagca gcagcaacag ccggcgcttc 10920  
tgccccgc ccagcagcag ccagccacta ccgcggcggc cgcctgagc ggagccggcg 10980  
ttcagtatga cctggccttg gaagagggcg aggggctggc gcgctgggg gcgtcgtcgc 11040  
cggagcggca cccgcgcgtg cagatgaaaa gggacgctcgc cgaggcctac gtgccaagc 11100  
agaaacctgt cagagacagc agcggcagc agcccagga gatcgcgcgc tcccgttcc 11160  
acgcggggcg ggagctcgcg cgcggcctg accgaaagc ggtcgtgagc gacgaggatt 11220  
tcgagcgga cgagctgacg gggatcagcc ccgcgcgcgc gcacgtggcc gcggccaacc 11280  
tggtcacggc gtacgagcag accgtgaaag aggagagcaa cttccaaaaa tccttcaaca 11340  
accacgtgcg cacgctgacg gcgcgcgagc aggtgaccct gggcctgatg cacctggtgg 11400  
acctgctgga ggcatcgtg cagaacccca cgagcaagcc gctgacggcg cagctgtttc 11460  
tgggtgtgca gcacagctcg gacaacgaga cgttcaggga ggcgctgctg aatataccg 11520  
agcccaggcg ccgctggctc ctggaactg tgaacattct gcagagcctc gtggtgcagc 11580  
agcgcgggct gccgctgtcc gagaagctg cggctatcaa cttctcgggt ctgagcctgg 11640  
gcaagtacta cgctaggaag atctacaaga ccccgtagc gccatagac aaggaggtga 11700  
agatcgacgg gttttacatg cgcagcagcc tgaagtgtc gacctgagc gacgatctg 11760  
gggtgtaccg caacgacagc atgcaccgcg cggtagcgc cagccgcgcg cgcgagctga 11820  
gcgaccagga gctgatgcac agcctgcagc gggccctgac cggggccggg accgagggg 11880  
agagctactt tgacatgggc gcggaactgc gctggcagcc cagccgcgcg gccttgggaag 11940

---

-continued

---

ctgccggcgg ttccccctac gtggaggagg tggacgatga ggaggaggag ggcgagtacc 12000  
tggagactg atggcgcgac cgtatttttg ctgatgacag caacagccac cgcctcctga 12060  
tcccgcgatg cgggcggcgc tgcagagcca gccgtccggc attaactcct cggacgattg 12120  
gaccacggcc atgcaacgca tcatggcgct gacgaccgc aatcccgaag cctttagaca 12180  
gcagcctcag gccaacggc tctcggccat cctggaggcc gtggtgccct cgcgctcga 12240  
ccccacgac gagaagggtc tggccatcgt gaacgcgctg gtggagaaca aggccatccg 12300  
cggcgacgag gccgggctgg tgtacaacgc gctgctggag cgcgtggccc gctacaacag 12360  
caccacgctg cagacgaacc tggaccgcat ggtgaccgac gtgcgagag cgggtgctgca 12420  
gcgagagcgg ttccaccgcg agtcgaacct gggctccatg gtggcgctga acgcttcct 12480  
gagcacgag cccgccaacg tccccgggg ccaggaggac tacaccaact tcatcagcgc 12540  
gctgcggtg atggtggccg aggtgcccc gagcaggtg taccagtcgg ggcgggacta 12600  
cttcttccag accagtcgcc agggcttgca gaccgtgaac ctgagccagg ctttcaagaa 12660  
cttgacggga ctgtggggcg tgcaggcccc ggtcggggac cgcgagcgg tgcgagcct 12720  
gctgacgccc aactcggccc tctgctgct gctggtggcg cccttcacgg acagcggcag 12780  
cgtgagccgc gactcgtacc tgggtacct gcttaacctg taccgagag ccatcgggca 12840  
ggcgacgctg gacgagcaga cctaccagga gatcaccac gtgagccgc cgtgggcca 12900  
ggaggacccc ggcaactcgg aggccacct gaacttcctg ctgaccaacc ggtcgcagaa 12960  
gatccccccc cagtacgccc tgagcaccga ggaggagcgc atcctgcgct acgtgcagca 13020  
gagcgtggg ctgttctcga tgcaggagg ggccacgccc agcgcgcgc tgcacatgac 13080  
cgcgcgcaac atggagccca gcatgtacgc tcgcaaccgc ccgttcatca ataagctgat 13140  
ggactacttg catcgggccc ccgccatgaa ctcggactac tttaccaacg ccatctgaa 13200  
ccccactg ctccccccc ccgggttcta cacgggcgag tacgacatgc ccgacccca 13260  
cgacgggttc ctgtgggac acgtggacag cagcgtgttc tcgcccgcgc ccgccaccac 13320  
cgtgtggaag aaagaggcgc gggaccgccc gccgtcctcg gcgctgtccg gtcgcccggg 13380  
tgctgcgcg gcggtgcctg aggcccccag ccccttccc agcctgcct tttcgtgaa 13440  
cagcgtgcgc agcagcagc tgggtcggct gacgcggccc cgcctgctgg gcgaggagga 13500  
gtacctgaac gactccttgt tgaggcccga gcgcgagaag aacttccca ataacgggat 13560  
agagagcctg gtggacaaga tgagccgctg gaagacgtac gcgacagc acagggacga 13620  
gccccgagct agcagcagc caggcaccgc tagacccag cgacacgaca ggcagcggg 13680  
tctggtgtgg gacgatgag attccgccga cgacagcagc gtgttgact tgggtgggag 13740  
tggtgtgtgt aaccgcttc ctcaettgccc ccccgatc gggcgcctga tgtaagaatc 13800  
tgaaaaaata aaaaacggta ctcaccaagg ccatggcgac cagcgtgcgt tcttctctgt 13860  
tgtttgtagt agtatgatga ggcgcgtgta cccggagggt cctcctcct cgtacgagag 13920  
cgtgatgac caggcgtgg cggcggcgat gcagccccg ctggaggcgc cttacgtgcc 13980  
ccccgggtac ctggcgccta cggagggcg gaacagcatt cgttactcg agctggcacc 14040  
cttgtagcat accaccgggt tgtacctggt ggacaacaag tcggcggaca tcgctcctgct 14100  
gaactaccag aacgaccaca gcaacttcct gaccaccgtg gtgcagaaca acgatttcac 14160  
ccccacggag gccagcacc agaccatcaa ctttgacgag cgtcgcgggt gggcgggcca 14220

---

-continued

---

gctgaaaacc atcatgcaca ccaacatgcc caacgtgaac gagttcatgt acagcaacaa 14280  
gttcaaggcg cgggtgatgg tctcgcgcaa gacccccaat ggggtcgcgg tggatgagaa 14340  
ttatgatggt agtcaggacg agctgactta cgagtgggtg gagtttgagc tgcccagagg 14400  
caacttctcg gtgaccatga ccacgatct gatgaacaac gccatcatcg acaactactt 14460  
ggcgggtggg cgtcagaacg ggggtctgga gagcgacatc ggcgtgaagt tcgacacgcg 14520  
caacttccgg ctgggctggg acccctgac cgagctgggtg atgccggggc tgtacaccaa 14580  
cgaggccttc cccccgaca tcgtcctgct gcccggtgc ggcgtggact tcaccgagag 14640  
ccgcctcagc aacctgctgg gcacccgcaa gcggcagccc tccaggagg gcttccagat 14700  
cctgtacgag gacctggagg ggggcaacat ccccgcgctc ttggatgtcg aagcctatga 14760  
gaaaagcaag gaggaggccg ccgcagcggc gaccgcagcc gtggccaccg cctctaccga 14820  
ggtgcggggc gataattttg ctagcggcgc ggagtggtcc gagcgggctg aaaccgaaag 14880  
taagatagtc atccagccgg tggagaagga cagcaaggac aggagctaca acgtgctcgc 14940  
ggacaagaaa aacaccgcct accgcagctg gtacctggcc tacaactacg gcgaccccga 15000  
gaaggcctg cgctcctgga cgctgctcac cacctcggac gtcacctgcg gcgtggagca 15060  
agtctactgg tcgctgccc acatgatgca agaccggctc accttccgct ccacgcgtca 15120  
agttagcaac taccgggtg tgggcgccga gctcctgccc gtctactcca agagcttctt 15180  
caacgagcag gccgtctact cgcagcagct gcgcgccttc acctcgtca cgcacgtctt 15240  
caaccgcttc cccgagaacc agatcctcgt ccgcccgcgc gcgcccacca ttaccaccgt 15300  
cagtgaaaac gttcctgctc tcacagatca cgggacctg ccgctgcgca gcagtatccg 15360  
gggagtccag cgcgtgaccg tcaactgacgc cagacggcgc acctgcccct acgtctacaa 15420  
ggccctgggc gtagtgcgc cgcgctcct ctcgagccgc accttctaaa aaatgtccat 15480  
tctcatctcg cccagtaata acaccggtg gggcctgcgc gcgcccagca agatgtacgg 15540  
aggcctcgc caacgctcca cgcaacacc cgtgcgcgtg cgcgggcaact tccgcgctcc 15600  
ctggggcgcc ctcaagggcc gcgtgcgctc gcgcaccacc gtcgacgacg tgatcgacca 15660  
ggtggtggcc gacgcgcgca actacacgcc gcgcgcgcg cccgcctcca ccgtggacgc 15720  
cgtcacgac agcgtggtg ccgatgcgc ccggtacgcc gcgccaaga gccggcggcg 15780  
gcgcatcgcc cggcggcacc ggagcaccgc gcctatgcgc gcggcgcgag ccttgctgcg 15840  
cagggcagg cgcacgggac gcagggccat gctcaggggc gccagacgc cggcctccgg 15900  
cagcagcagc gccggcagga cccgcagacg cgcggccacg gcggcggcgg cggccatcgc 15960  
cagcatgtcc cgcgcgcgc gcggcaact gtactgggtg cgcgacgccc ccaccggtgt 16020  
gcgctgccc gtgcgacccc gccccctcg cacttgaaga tgcgtacttc gcgatgttga 16080  
tgtgtccag cggcagagg gatgtccaag cgcaataca aggaagagat gctccaggtc 16140  
atcgcgcctg agatctacgg ccccgcggtg aaggaggaaa gaaagccccg caaactgaag 16200  
cgggtcaaaa aggacaaaa ggaggaggaa gatgtggacg gactggtgga gtttgtgcgc 16260  
gagttcgcgc cccggcggcg cgtgcagtgg cgcgggcgga aagtgaacc ggtgctgcg 16320  
cccggaccca cgggtgtctt cacgcccgc gagcgttccg gctccgctc caagcgtcc 16380  
tacgacgagg tgtacgggga cgaggacatc ctcgagcagg cgtcgcgagc tctgggcgag 16440  
tttgcttacg gcaagcgcag ccgcccgcgc cccttgaag aggaggcggg gtccatccc 16500

---

-continued

---

ctggaccacg gcaacccccac gccgagcctg aagccggtga ccctgcagca ggtgctgccg 16560  
agcgcggcgc cgcgcggggg cttcaagcgc gagggcgcg aggatctgta cccgaccatg 16620  
cagctgatgg tgcccaagcg ccagaagctg gaggacgtgc tggagcacaat gaaggtggac 16680  
cccgaggtgc agcccaggt caaggtgcgg cccatcaagc aggtggcccc gggcctgggc 16740  
gtgcagaccg tggacatcaa gatccccacg gagcccatgg aaacgcagac cgagcccgtg 16800  
aagcccagca ccagcaccat ggaggtgcag acggatccct ggatgccggc gccggcttcc 16860  
accactcgcc gaagacgcaa gtacggcgcg gccagcctgc tgatgcctaa ctacgcgctg 16920  
cctccttcca tcatccccac gccgggttac cgcggcacgc gcttctaccg cggctacacc 16980  
agcagccgcc gcaagaccac caccgcgcgc cgcctcgtc gcacccgccg cagcagcacc 17040  
gcgacttccg ccgcccctt ggtgcggaga gtgtaccgca gcgggcgcga gcctctgacc 17100  
ctgccgcgcg cgcgctacca cccgagcctc gccatttaac tctgccgtcg cctcctactt 17160  
gcagatatgg ccctcacatg ccgcctccgc gtccccatta cgggctaacc aggaagaaag 17220  
ccgcgcgcta gaaggtgac ggggaacggg ctgcgtgcc atcaccaccg gcggcggcgc 17280  
gccatcagca agcgggtggg gggaggcttc ctgccgcgc tgatccccat catcgccgcg 17340  
gcgatcgggg cgatccccg catagcttcc gtggcggtgc aggcctctca gcgccactga 17400  
gacacagctt ggaatattg taataaaaa atggactgac gctcctggtc ctgtgatgtg 17460  
tgtttttaga tggaaagacat caatttttcg tccctggcac cgcgacacgg cacgcggccg 17520  
tttatgggca cctggagcga catcggcaac agccaactga acgggggcgc cttcaattgg 17580  
agcagtctct ggagcgggt taagaatttc ggggccacgc tcaaaccta tggcaacaag 17640  
gcgtggaaca gcagcacag gcagcgctg agggaaaagc tgaagagca gaacttccag 17700  
cagaaggtgg tcgatggcct gccctcgggc atcaacgggg tggaggacct ggccaaccag 17760  
gccgtgcaga aacagatcaa cagccgcctg gacgcggtcc cgcgccggg gtcctgtggag 17820  
atgccccagc tggaggagga gctgcctccc ctggacaagc gcggcgacaa gcgaccgctg 17880  
cccgacgcgg aggagacgct gctgacgcac acggacgagc cgcccccgta cgaggaggcg 17940  
gtgaaactgg gtctgcccac cacgcggccc gtggcgctc tggccaccgg ggtgctgaaa 18000  
cccagcagca gcagccagcc cgcgacctg gacttgctc cgcctgctc ccgccctcc 18060  
acagtggcta agcccctgcc gccggtggcc gtcgctcgc gcgcccccg aggcgcgcc 18120  
caggcgaact ggcagagcac tctgaacagc atcgtgggtc tgggagtga gagtgtgaag 18180  
cgcgccgct gctattaaaa gacactgtag cgcttaactt gcttgtctgt gtgtatatgt 18240  
atgtccgccg accagaagga ggaagagcg cgtcgccgag ttgcaagatg gccaccccat 18300  
cgatgctgcc ccagtggcg tacatgcaca tcgccggaca ggacgcttcg gagtacctga 18360  
gtccgggtct ggtgcagttc gcccgcgcca cagacaccta cttcagtctg ggaacaagt 18420  
ttaggaacct cacggtggcg cccacgcacg atgtgaccac cgaccgcagc cagcgctga 18480  
cgctgcgctt cgtgccctg gaccgcgag acaacaccta ctcgtacaaa gtgcgctaca 18540  
cgctggccgt gggcgacaac cgcgtgctg acatggccag cacctacttt gacatccgcg 18600  
gcgtgctgga tcgggggccc agcttcaaac cctactccgg caccgcctac aacagcctgg 18660  
ctcccaagg agcgcacca acttgccagt ggacatataa agctgggtgat actgatacag 18720  
aaaaaaccta tacatatgga aatgcacctg tgcaaggcat tagcattaca aaggatggtg 18780

---

-continued

---

ttcaacttgg aactgacagc gatggtcagg caatctatgc agacgaaact tatcaaccag 18840  
agcctcaagt gggatgatgct gaatggcatg acatcactgg tactgatgaa aaatatggag 18900  
gcagagctct taagcctgac accaaaatga agccttgcta tggttctttt gccaaagccta 18960  
ccaataaaga aggaggccag gcaaatgtga aaaccgaaac aggcggtacc aaagaatatg 19020  
acattgacat ggcattcttc gataatcgaa gtgcagctgc cgccggccta gccccagaaa 19080  
ttgttttgta tactgagaat gtggatctgg aaactccaga taccatatt gtatacaagg 19140  
caggtagaca tgacagtagc tcttctatca atttgggtca gcagtcctatg cccaacagac 19200  
ccaactacat tggcttcaga gacaacttta tcggtctgat gtactacaac agcactggca 19260  
atatgggtgt actggctgga caggcctccc agctgaatgc tgtggtggac ttgcaggaca 19320  
gaaacaccga actgtcctac cagctcttgc ttgactctct ggtgacaga accaggtatt 19380  
tcagtatgtg gaatcaggcg gtggacagtt atgaccccga tgtgcgcatt attgaaaatc 19440  
acggtgtgga ggatgaactt cctaactatt gcttccccct ggatgctgtg ggtagaactg 19500  
atacttacca gggaattaag gccaatggtg ataataaac cacctggacc aaagatgata 19560  
ctgttaatga tgctaataaa ttgggcaagg gcaatccttt cgccatggag atcaacatcc 19620  
aggccaacct gtggcggaac ttctctacg cgaacgtggc gctgtacctg cccgactcct 19680  
acaagtacac gccggccaac atcacgctgc ccaccaaac caacacctac gattacatga 19740  
acggcccgct ggtggcgccc tcgctggtgg acgctacat caacatcggg gcgcgctggt 19800  
cgctggacc ccatggacaac gtcaaccctt tcaaccacca ccgcaacgcg gccctgcgat 19860  
accgtccat gctcctgggc aacggcgct acgtgccctt ccacatccag gtgccccaaa 19920  
agtttttcgc catcaagagc ctctctctcc tgcccgggtc ctacacctac gagtggaaact 19980  
tccgcaagga cgtcaacatg atcctgcaga gctcctcgg caacgacctg cgcacggacg 20040  
gggctccat cgccttcacc agcatcaacc tctacgccac cttcttcccc atggcgcaaca 20100  
acaccgcctc cagcctcgag gccatgctgc gcaacgacac caacgaccag tccttcaacg 20160  
actacctctc ggcgccaac atgctctacc ccatcccggc caacgccacc aacgtgcca 20220  
tctccatccc ctgcgcaac tgggcgctt tccgcgctg gtccttcacg cgcctcaaga 20280  
cccgcgagac gcctcgcctc ggtccgggt togacccta cttcgtctac tcgggctcca 20340  
tcccctacct cgacggcacc ttctacctca accacacctt caagaaggtc tccatcacct 20400  
tcgactcctc cgtcagctgg cccggcaacg accgctcct gacgccaac gaggtcgaaa 20460  
tcaagcgac cgtcgacgga gagggtaca acgtggcca gtgcaacatg accaaggact 20520  
ggttctggt ccagatgctg gccactaca acatcggtta ccagggttc tacgtgcccg 20580  
agggctaca ggaccgatg tactccttct tccgcaactt ccagccatg agcgcgagg 20640  
tcgtggacga ggtcaactac aaggactacc aggcctcac cctggcctac cagcacaaca 20700  
actcgggctt cgtcggctac ctgcgcca ccatgcgcca gggccagccc taccgcca 20760  
actacccta cccgctcctc ggcaagagcg ccgctgcag cgtcaccag aaaaagtcc 20820  
tctcggacc ggtcatgtg cgcctcccct tctccagcaa cttcatgtcc atggcgcg 20880  
tcaccgacct cgccagaac atgctctac ccaactccgc ccacgcgcta gacatgaatt 20940  
tcgaagtca cccatggat gactccacc ttctctatgt tgtcttcaa gtcttcgacg 21000  
tcgtccgagt gcaccagccc caccggcgct tcatcgaggc cgtctacctg cgcacgccct 21060

---

-continued

---

tctcggcccg caacgccacc acctaagcct cttgcttctt gcaagatgac ggctcgcgcg 21120  
ggctccggcg agcaggagct cagggccatc ctcccgacc tgggctgcgg gccctgcttc 21180  
ctgggcacct tcgacaagcg cttcccgga ttcatggccc cgcacaagct ggctcgcgcc 21240  
atcgtcaaca cggccggccg cgagaccggg ggcgagcact ggctggcctt cgctcggaac 21300  
ccgcgctccc acacctgcta cctcttcgac cccttcgggt tctcggacga gcgctcaag 21360  
cagatctacc agttcgagta cgaggcctg ctgcgtcgca gcgccctggc caccgaggac 21420  
cgctgcgtca ccttgaaaa gtccaccag accgtgcagg gtccgcgctc ggccgctgc 21480  
gggctcttct gctgcatgtt cctgcacgcc ttcgtgcact ggcccgaccg ccccatggac 21540  
aagaaccca ccatgaactt gctgacgggg gtgccaacg gcatgctcca gtcgcccag 21600  
gtggaaccca cctgcgccc caaccaggag gcgctctacc gcttccctca cgcctctcc 21660  
gcctactttc gctcccaccg cgcgcgcatc gagaaggcca ccgcttcga ccgcatgaat 21720  
caagacatgt aatccgggtg gtgtatgtga atgctttatt catcataata aacagcat 21780  
gtttatgcca cttctctga ggctctgact ttatttagaa atcgaagggg ttctgcccgc 21840  
tctcggcatg gcccgggcg agggatacgt tgcggaactg gtacttgggc agcacttga 21900  
actcgggat cagcagcttc ggcacgggga ggtcgggaa cgagtcgctc cacagcttc 21960  
gcgtgagttg cagggcgccc agcaggtcgg gcgcgagat cttgaaatcg cagttgggac 22020  
ccgcgttctg cgcgcgagag ttacggtaca cggggttgca gcaactggaac accatcaggg 22080  
ccgggtgctt cagctcgcg agcaccgtcg cgtcgggtgat gccctccacg tccagatcct 22140  
cggcgttggc catcccgaag ggggtcatct tgcaggtctg ccgcccctg ctgggcacgc 22200  
agccgggctt gtggtgcaa tcgcagtgca ggggatcag catcatctgg gcctgctcgg 22260  
agctcatgcc cgggtacatg gccttcatga aagcctccag ctggcggaag gcctgctcgc 22320  
ccttgccgccc ctccggtgaag aagaccccgc aggacttctt agagaactgg ttggtggcgc 22380  
agccagcgtc gtgcacgca cagcgcgctg cgttgttggc cagctgcacc acgctcgcgc 22440  
cccagcgggt ctgggtgatc ttggcccggc cggggttctc cttcagcgcg cgctgcccgt 22500  
tctcgtcgc cacatccatc tcgatcgtgt gctcctctg gatcatcacg gtcocgtgca 22560  
ggcaccgca cttgccctcg gcctcgtgc acccgtgcag ccacagcgcg cagccggtgc 22620  
tctcccagtt cttgtggcg atctgggagt gcgagtgca gaagccctgc aggaagcggc 22680  
ccatcatcgt ggtcagggtc ttgttctcgg tgaaggtcag cggaatgccg cgtgctcct 22740  
cgttcacata caggtggcag ataccggcgt acacctcgc ctgctcgggc atcagctgga 22800  
aggcggactt caggtcgcgc tccacgcggt accggtccat cagcagcgtc atcactcca 22860  
tgcccttctc ccaggccgaa acgatcggca ggctcagggg gttcttcacc gttgtcatct 22920  
tagtcggcgc cgcgaagtc agggggtcgt tctcgtccag ggtctcaaac actcgttgc 22980  
cgtccttctc ggtgatgcgc acgggggaa agctgaagcc cacggccgccc agctcctcct 23040  
cggcctgcct ttcgtcctcg ctgctcctgg tgatgtcttg caaaggcaca tgcctggctc 23100  
tgccggggtt ctttttgggc ggcagaggcg gcggcgaga cgtcgtgggc gagcgcgagt 23160  
tctcgtcac cacgactatt tcttctcctt ggccgtcgtc cgagaccacg cggcggtagg 23220  
catgcctctt ctggggcaga ggcggaggcg acgggctctc gcggttcggc gggcggctgg 23280  
cagagccctt tccgcgttcg ggggtgcgct cctggcggcg ctgctctgac tgaactctc 23340

---

-continued

---

cgcgccggc cattgtgttc tcctagggag caagcatgga gactcagcca tcgtcgccaa 23400  
catcgccate tgccccgcc gccgcccagc agaaccagca gcagcagaat gaaagcttaa 23460  
ccgccccgcc gccagcccc acctccgagc ccgagcccc agacatgcaa gagatggagg 23520  
aatccatcga gattgacctg ggctacgtga cgcccggga gcacgaggag gagctggcag 23580  
cgcgcttttc agccccgaa gagaaccacc aagagcagcc agagcaggaa gcagagagcg 23640  
agcagaacca ggctgggctc gagcatggcg actacctgag cggggcagag gacgtgctca 23700  
tcaagcatct ggcccccaa tgcacatcgc tcaaggacgc gctgctcgac cgcgccgagg 23760  
tgccccctag cgtggcggag ctccagccgc cctacgagcg caacctcttc tcgcccgcg 23820  
tgcccccaa gccccagccc aacggcaact gcgagccaa ccgcccctc aacttctacc 23880  
cggctctcgc ggtgcccag gccctggcca cctaccacct ctttttcaag aaccaaagga 23940  
tccccgtctc ctgccggcc aaccgcaccc gcgcccagc cctgctcaac ctgggccccg 24000  
gcgcccgcct acctgatata gcctccttgg aagaggttcc caagatcttc gaggtcttg 24060  
gcagcgacga gactcgggcc gcgaacgctc tgcaaggaa cggagaggag catgagcacc 24120  
acagcgccct ggtggagttg gaaggcgaca acgcccgcct ggcggtcctc aagcgacgg 24180  
tcgagctgac ccacttcgcc taccggcgc tcaacctgcc cccaaggtc atgagcgccg 24240  
tcatggacca ggtgctcctc aagcgcgct cgcccctctc ggaggaggag atgcaggacc 24300  
ccgagagctc ggacgagggc aagcccgtg tcagcgacga gcagctggcg cgtggcttg 24360  
gagcgagtag cccccccag agcctggaag agcggcgcaa gctcatgatg gccgtggtcc 24420  
tggtgacctg ggagctggag tgtctgcgcc gcttctctgc cgacgaggag acctgcgca 24480  
aggtcgagga gaacctgcac tacctcttca gacacgggtt cgtgcgccag gcctgcaaga 24540  
tctccaacgt ggagctgacc aacctggtct cctacatggg catcctgcac gagaaccgcc 24600  
tggggcagaa cgtgctgcac accaccctgc gcggggaggc ccgcccgcac tacatccgg 24660  
actggtctca cctgtacctc tgccacacct ggcagacggg catgggcgtg tggcagcagt 24720  
gcctggagga gcagaacctg aaagagctct gcaagctcct gcagaagaac ctcaaggccc 24780  
tgtggaccgg gttcgacgag cgcaccaccg ccgcccagct ggcagacctc atcttccccg 24840  
agcgcctgag gctgacgctg cgcaacgggc tgcccgactt tatgagccaa agcatgttg 24900  
aaaactttcg ctctttcctc ctccaacgct ccgggacctc gcccgccacc tgctccgcg 24960  
tgccctcgga ctctgtgccc ctgaccttcc gcgagtgcc cccgccgctc tggagcact 25020  
gctacctgct gcgctggccc aactaactgg cctaccactc ggacgtgatc gaggacgtca 25080  
gcggcgaggg cctgctogag tgccactgcc gctgcaacct ctgacgccc caccgctccc 25140  
tgccctgcaa cccccagctg ctgagcgaga cccagatcat cggcaccttc gaggttgcaag 25200  
gccccggcga gggcaagggg ggtctgaaac tcacccggg gctgtggacc tcggcctact 25260  
tgcgcaagtt cgtgcccag gactaccatc ccttcgagat caggttctac gaggaccaat 25320  
cccagccgcc caaggccgag ctgtcggcct gcgtcatcac ccagggggcc atcttgccc 25380  
aattgcaagc catccagaaa tcccgcgaag aatttctgct gaaaaagggc cacggggtct 25440  
acttgacc ccagaccgga gaggagctca accccagctt cccccagat gccccgagga 25500  
agcagcaaga agctgaaagt ggagctgccc ccgcccggg aggatgtgga ggaagactgg 25560  
gagagcagtc aggcagagga ggaggagatg gaagactggg acagcactca ggcagaggag 25620



---

-continued

---

gacagcctgc aagacagtct ggaggaggaa gacgaggtgg aggaggcaga ggaagaagca 25680  
gccgccgcca gaccgtcgtc ctccggcggag gaggagaaag caagcagcac ggataccatc 25740  
tccgctccgg gtccgggtcg cggcggccgg gcccacagta gatgggacga gaccggggcg 25800  
ttccgaaacc ccaccacca gaccggtaag aaggagcggc agggatacaa gtccctggcg 25860  
gggcacaaaa acgccatcgt ctccctgctg caagcctgcg ggggcaacat ctccctcacc 25920  
cggcgtacc tgctcttcca ccgcggggtg aacttcccc gcaacatctt gcattactac 25980  
cgtcacctcc acagccccta ctactgttcc caagaagagg cagaaacca gcagcagcag 26040  
cagcagcaga aaaccagcgg cagcagctag aaaatccaca gcggcggcag gtggactgag 26100  
gatcgcggcg aacgagccgg cgcagaccgg ggagctgagg aaccggatct tccccacct 26160  
ctatgccatc ttccagcaga gtcgggggca agagcaggaa ctgaaagtca agaaccgttc 26220  
tgtcgtctcg ctccccgca gttgtctgta tcacaagagc gaagaccaac ttcagcgcac 26280  
tctcagggac gccgaggctc tcttcaacaa gtactgcgcg ctactctta aagagtagcc 26340  
cgcgccgccc cacacacgga aaaaggcggg aattacgtca ccacctgcgc ccttcgccc 26400  
accatcatca tgagcaaaga gattcccacg ccttacatgt ggagctacca gcccagatg 26460  
ggcctggccg ccggcggcgc ccaggactac tcccccgca tgaactggct cagtgcgggg 26520  
cccgcgatga tctcacgggt gaatgacatc cgcgccacc gaaaccagat actcctagaa 26580  
cagtcagcga tcaccgccac gccccgccat caccttaatc cgcgtaattg gcccgccgc 26640  
ctggtgtacc aggaaattcc ccagcccacg accgtactac ttccgcgaga cggccaggcc 26700  
gaagtccagc tgactaaact aggtgtccag ctggcggcg gcgccccct gtgtcgtcac 26760  
cgccccgctc aggtataaa cgggtgtgtg atccgaggca gaggcacaca gctcaacgac 26820  
gaggtgtgta gctcttcgct gggctcgcga cctgacggag tcttccaact cgcggatcg 26880  
gggagatctt ccttcacgcc tcgtcaggcc gtcctgactt tggagagttc gtcctcgag 26940  
ccccgctcgg gtggcatcgg cactctccag ttctgtggagg agttcactcc ctccgtctac 27000  
ttcaaccctc tctccggctc ccccgccac taccggacg agttcatccc gaacttcgac 27060  
gccatcagcg agtcggtgga cggtacgat tgaatgtccc atggtggcgc ggtgacctc 27120  
gctcggcttc gacacctgga ccactgccgc cgttccgct gcttcgctcg ggatctcgcc 27180  
gagtttgctc actttgagct gcccgaggag caccctcagg gcccgccca cggagtgcgg 27240  
atcgtcgtcg aaggggtctc cgaactccac ctgcttcgga tcttcagcca gcgtccgatc 27300  
ctggccgagc gcgagcaagg acagaccctt ctgaccctgt actgcatctg caaccacccc 27360  
ggcctgcatg aaagtctttg ttgtctgctg tgtactgagt ataataaaag ctgagatcag 27420  
cgactactcc ggacttccgt gtgttctcgc tatcaaccag tccctgttct tcaccgggaa 27480  
cgagaccgag ctccagctcc agtgaagcc ccacaagaag tacctcaoct ggctgttcca 27540  
gggctctccg atcggcgttg tcaaccactg cgacaacgac ggagtcctgc tgagcggccc 27600  
tgccaaacct actttttcca cccgcagaag caagctccag ctcttccaac ccttcctccc 27660  
cgggacctat cagtcgctct cgggaacctg coactcacacc ttccacctga tccogaatac 27720  
cacagcgtcg ctccccgcta ctaacaacca aactaccac caacgccacc gtcgcgacct 27780  
ttcctctggg tctaatacca ctaccggagg tgagctccga ggtcgaccaa cctctgggat 27840  
ttactacggc ccctgggagg tggtagggtt aatagcgcga ggcctagttg cgggtgggct 27900

---

-continued

---

tttgctctc tgctacctat acctcccttg ctgttcgtac ttagtggtgc tgtgttgctg 27960  
gtttaagaaa tggggaagat caccctagtg agctgcggtg tgctgggtggc ggtgggtgctt 28020  
tcgattgtgg gactgggocgg cgcggtgta gtgaaggaga aggccgatcc ctgcttgcat 28080  
ttcaatcccc acaaatgcca gctgagtttt cagccccgatg gcaatcggtg cgcggtgctg 28140  
atcaagtgcg gatgggaatg cgagaacgtg agaatcgagt acaataacaa gactcggaac 28200  
aatactctcg cgtccgtgtg gcagccocgg gaccccgagt ggtacaccgt ctctgtcccc 28260  
ggtgctgacg gtccccccg caccgtgaat aatactttca tttttgcgca catgtgcgac 28320  
acggctatgt ggatgagcaa gcagtacgat atgtggcccc ccacgaagga gaacatcgtg 28380  
gtcttctcca tcgcttacag cgtgtgcacg cgcctaata cgcctatcgt gtgcctgagc 28440  
attcacatgc tcatcgctat tcgccccaga aataatgccg aaaagaaaa acagccataa 28500  
cacgtttttt cacacacctt tttcagacca tggcctctgt taaatttttg cttttatttg 28560  
ccagtctcat tgccgtcatt catggaatga gtaatgagaa aattactatt tacactggca 28620  
cfaatcacac attgaaaggt ccagaaaaag ccacagaagt ttcatggtat tgttatttta 28680  
atgaatcaga tgtatctact gaactctgtg gaaacaataa caaaaaaat gagagcatta 28740  
ctctcatcaa gtttcaatgt ggatctgact taaccctaata taacatcact agagactatg 28800  
taggtatgta ttatggaact acagcaggca tttcggacat ggaattttat caagtttctg 28860  
tgtctgaacc caccacgcct agaatgacca caaccacaaa aactacacct gttaccacta 28920  
tacagctcac taccaatggc tttcttgcca tgcttcaagt ggctgaaaat agcaccagca 28980  
ttcaaccacc cccaccaggt gaggaattc ccagatccat gattggcatt attggtgctg 29040  
tagtggtgtg catgttgatc atcgccttgt gcatggtgta ctatgccttc tgctacagaa 29100  
agcacagact gaacgacaag ctggaacact tactaagtgt tgaattttaa ttttttagaa 29160  
ccatgaagat cctagccctt ttagtttttt ctatcattac ctctgctcta tgcaattctg 29220  
acaatgagga cgttactgtc gttgtoggat caaattatac actaaaaggt ccagcaaaag 29280  
gtatgctttc gtggatttgt tggttcggaa ctgacgagca acagacagaa ctttgcaatg 29340  
ctcaaaaag caaaacctca aattcaaaa tctctaatta tcaatgcaat ggcactgact 29400  
tagtattgct caatgtcacg aaagcatatg ctggcagtta cacctgcoct ggagatgatg 29460  
ccgacaatat gattttttac aaagtggaag tggttgatcc cactactcca ccgcccacca 29520  
ccacaactac tcataccaca cacacagaac aaacaccaga ggcagcagaa gcagagttgg 29580  
ccttcaggt tcacggagat tcctttgctg tcaatacccc tacaccgat cagcgggtgc 29640  
cggggctgct cgtcagcggc attgtcggtg tgctttcggg attagcagtc ataatcatct 29700  
gcatgttcat ttttgcttgc tgctatagaa ggctttaccg acaaaaatca gaccactgc 29760  
tgaacctcta tgtttaatth tttccagagc catgaaggca gttagcgtc tagttttttg 29820  
ttctttgatt ggcattgttt ttagtgctgg gtttttgaat aatcttacca tttatgaag 29880  
tgagaatgcc actctagtgg gcatcagtg tcaaatgctc agctggctaa aataccatct 29940  
agatgggtgg aaagacattt gcgattgaa tgtcactgtg tatacatgta atggagttaa 30000  
cctcaccatt actaatgcca cccaagatca gaatggtagg ttttaaggcc agagtttcac 30060  
tagaaaatg gggatgaaat ccataacat gtttatctat gacgtcactg tcatcagaaa 30120  
tgagactgcc accaccacac agatgcccac tacacacagt tctaccacta ctaccatgca 30180

-continued

---

|            |            |             |             |            |            |       |
|------------|------------|-------------|-------------|------------|------------|-------|
| aaccacacag | acaaccacta | catcaactca  | gcatatgacc  | accactacag | cagcaaagcc | 30240 |
| aagtagtgca | gcgctcagc  | cccaggcttt  | ggctttgaaa  | gctgcacaac | ctagtacaac | 30300 |
| tactaggacc | aatgagcaga | ctactgaatt  | tttgtccact  | gtcgagagcc | acaccacagc | 30360 |
| tacctcagc  | gccttctcta | gcaccgcaa   | tctctcctcg  | cttctctcta | caccaatcag | 30420 |
| tcccgcact  | actcccacc  | cagctcttct  | ccccactccc  | ctgaagcaaa | ctgaggacag | 30480 |
| cgcatgcaa  | tggcagatca | ccctgctcat  | tgtgatcggg  | ttggtcatcc | tggccgtgtt | 30540 |
| gctctactac | atcttctgcc | gcccattcc   | caacgcgcac  | cgcaaaccgg | cctacaagcc | 30600 |
| catcgttate | gggcagccgg | agccgcttca  | ggtggaaggg  | ggtctaagga | atcttctctt | 30660 |
| ctcttttaca | gtatggtgat | tgaactatga  | ttcctagaca  | attcttgatc | actattctta | 30720 |
| tctgcctcct | ccaagtctgt | gccaccctcg  | ctctggtggc  | caacgccagt | ccagactgta | 30780 |
| ttgggccctt | cgctcctac  | gtgctctttg  | ccttcatcac  | ctgcatctgc | tgctgtagca | 30840 |
| tagtctgcct | gcttatcacc | ttcttccagt  | tcattgactg  | gatctttgtg | cgcatcgcct | 30900 |
| acctgcgcca | ccacccccag | taccgcgacc  | agcgagtggc  | gcgctgctc  | aggctcctct | 30960 |
| gataagcatg | cggtctctgc | tacttctcgc  | gcttctgctg  | ttagtgtctc | cccgcccggt | 31020 |
| cgacccccgg | tccccactc  | agtccccga   | agaggctccg  | aaatgcaaat | tccaagaacc | 31080 |
| ctggaaatc  | ctcaaatgct | accgcaaaa   | atcagacatg  | cttcccagct | ggatcatgat | 31140 |
| catgggatc  | gtgaacatc  | tggctgcac   | cctcatctcc  | tttgtgattt | accctgctt  | 31200 |
| tgactttggt | tggaactcgc | cagagggcgt  | ctatctcccg  | cctgaacctg | acacaccacc | 31260 |
| acagcaacct | caggcacacg | cactaccacc  | accacagcct  | aggccacaat | acatgcccat | 31320 |
| attagactat | gaggccgagc | cacagcgacc  | catgctcccc  | gctattagtt | acttcaatct | 31380 |
| aaccggcgga | gatgactgac | ccactggcca  | acaacaacgt  | caacgacctt | ctcctggaca | 31440 |
| tggacggccg | cgctcggag  | cagcgactcg  | cccacttctg  | cattcgccag | cagcaggaga | 31500 |
| gagccgtcaa | ggagctcgag | gacggcatag  | ccatccacca  | gtgcaagaaa | ggcatcttct | 31560 |
| gcctggtgaa | acaggccaag | atctcctacg  | aggtcacccc  | gaccgacctt | cgctctcct  | 31620 |
| acgagctcct | gcagcagcgc | cagaagttca  | cctgctgggt  | cggagtcaac | ccatcgtca  | 31680 |
| tcaccagca  | gtcggggcat | accaaggggt  | gcattccactg | ctcctgcgac | tccccgact  | 31740 |
| gctgccacac | tctgatcaag | accctctcgc  | gcctccgca   | cctcctcccc | atgaactaat | 31800 |
| caccccccta | tccagtgaaa | taaataatcat | attgatgatg  | atttaaataa | aaaataatca | 31860 |
| tttgatttga | aataaagata | caatcatatt  | gatgatttga  | gttttaaaaa | ataaagaatc | 31920 |
| acttacttga | aatctgatac | caggtctctg  | tccatgtttt  | ctgccaacac | cacctcactc | 31980 |
| ccctcttccc | agctctggtg | ctgcagaccc  | cggcggtctg  | caaacttctt | ccacacgctg | 32040 |
| aaggggatgt | caaatctctc | ctgtcctcca  | atcttcattt  | tatcttctat | cagatgtcca | 32100 |
| aaaagcgcgt | ccgggtggat | gatgactctg  | acccgctcta  | cccctacgat | gcagacaacg | 32160 |
| caccgacctg | gcccctcctc | aacccccctt  | tgtctcttct  | agatggattc | caagagaagc | 32220 |
| ccctgggggt | gctgtccctg | cgactggctg  | acccgctcac  | caccaagaac | ggggaaatca | 32280 |
| ccctcaagct | gggagagggg | gtggacctcg  | actcctcggg  | aaaactcatc | tccaacacgg | 32340 |
| ccaccaaggc | cgccgcccct | ctcagttttt  | ccaacaacac  | catttccctt | aacatggata | 32400 |
| cccccttcta | taccaagatg | ggaaaattat  | ccttacaagt  | ttctccaccg | ttaaacatat | 32460 |

---

-continued

---

taaaatcaac cattctgaac acattagctg tagcttatgg atcaggttta ggactgagtg 32520  
gtggcactgc tcttgcaagta cagttggcct ctccactcac ttttgatgaa aaaggaaata 32580  
ttaaattaa cctagccagt ggtccattaa cagttgatgc aagtcgactt agtatcaact 32640  
gcaaaagagg ggtcactgtc actacctcag gagatgcaat tgaagcaac ataagctggc 32700  
ctaaaggat aagatttgaa ggtaatggca tagctgcaaa cattggcaga ggattggaat 32760  
ttggaaccac tagtacagag actgatgtca cagatgcata cccaattcaa gttaaattgg 32820  
gtactggcct tacctttgac agtacaggcg ccattgttgc ttggaacaaa gaggatgata 32880  
aacttacatt atggaccaca gccgaccctc cgccaattg caaaatatac tctgaaaaag 32940  
atgcaaaact cacactttgc ttgacaaagt gtggaagtca aattctgggt actgtgactg 33000  
tattggcagt gaataatgga agtctcaacc caatcacaaa cacagtaagc actgcactcg 33060  
tctcctcaa gtttgatgca agtggagttt tgctaagcag ctccacatta gacaaagaat 33120  
attggaactt cagaaagga gatgttacac ctgctgagcc ctataactaa gctataggtt 33180  
ttatgcctaa cataaaggcc tatcctaaaa acacatctgc agcttcaaaa agccatattg 33240  
tcagtcaagt ttatctcaat ggggatgag ccaaaccact gatgctgatt attactttta 33300  
atgaaactga ggatgcaact tgcacctaca gtatcacttt tcaatggaaa tgggatagta 33360  
ctaagtacac aggtgaaaca cttgtacca gctccttca cttctcctac atcgccaag 33420  
aatgaacact gtatcccacc ctgcatgcca acccttccca cccactctg tctatggaaa 33480  
aaactctgaa gcacaaaata aaataaagtt caagtgtttt attgattcaa cagttttaca 33540  
ggattcgagc agttattttt cctccaccct cccaggacat ggaatacacc accctctccc 33600  
cccgcacagc cttgaacatc tgaatgcat tggatgatgga catgcttttg gtctccacgt 33660  
tccacacagt ttcagagcga gccagctctg ggtcggtcag ggagatgaaa ccctccgggc 33720  
actcccgcct ctgcaacctc agctcaaca gctgaggatt gtctcggtg gtcgggatca 33780  
cggttatctg gaagaagcag aagagcggcg gtgggaatca tagtccgca acgggatcgg 33840  
ccggtggtgt cgcacagcg cccgcagcag tgcctgcccgc cgcctctccg tcaagctgct 33900  
gtcaggggg tccgggtcca gggactccct cagcatgatg cccacggccc tcagcatcag 33960  
tcgtctggtg cggcggcgcg agcagcgcct ggggatctcg ctcaggctcg tgcagtacgt 34020  
gcaacacagc accaccaggt tgttcaacag tccatagtcc aacacgctcc agccgaaact 34080  
catcgcggga aggatgctac ccacgtggcc gtcgtaccag atcctcaggt aaatcaagtg 34140  
gcgctccctc cagaacacgc tgcacacgta catgatctcc ttgggcatgt ggcggttcac 34200  
cacctcccgg taccacatca ccctctggtt gaacatgca ccccgatga tcctgcggaa 34260  
ccacagggcc agcaccgccc cgcctcccat gcagcgaaga gaccccggt cccggcaatg 34320  
gcaatggagg acccaccgct cgtaccctg gatcatctgg gagctgaaca agtctatggt 34380  
ggcacagcac aggcataatc tcatgcatct cttcagcact ctcagctcct cgggggtcaa 34440  
aaccatattc cagggcagcg ggaactcttg caggacagcg aaccccgag aacagggcaa 34500  
tcctcgca taacttatct tgtgatgga cagggatcag caatcaggca gcaaccgggtg 34560  
atcctccacc agagaagcgc gggctctggt ctcctcacag cgtggtaag gggccggccg 34620  
atacgggtga tggcgggagc cggctgatcg tgttcgagc cgtgtcatga tgcagttgct 34680  
ttcggacatt ttcgtacttg ctgtagcaga acctggtccg ggcgctgac accgatcgcc 34740

-continued

---

```

ggcggcggtc cggcgcttg gaacgctcgg tgttgaaatt gtaaacacgc cactctctca 34800
gaccgtgcag cagatctagc gcctcaggag tgatgaagat cccatcatgc ctgatagctc 34860
tgatcacatc gaccaccgtg gaatgggcca gaccagcca gatgatgcaa tttgtgtggg 34920
tttcggtgac ggcgggggag ggaagaacag gaagaacat gattaacttt taatccaaac 34980
ggtctcggag cacttcaaaa tgaaggtcgc ggagatggca cctctcgcgc cgcctgtgtt 35040
ggtgaaaaat aacagccagg tcaaaggta tacggttctc gagatgttcc acggtggctt 35100
ccagcaaagc ctccacgcgc acatccagaa acaagacaat agcgaagcg ggagggttct 35160
ctaattcctc aatcatcatg ttacactcct gcaccatccc cagataattt tcatttttcc 35220
agccttgaat gattcgaact agttcctgag gtaaatcaa gccagccatg ataaagagct 35280
cgcgcagagc gccctccacc ggcattctta agcacacct cataattcca agatattctg 35340
ctcctgggtc acctgcagca gattgacaag cggaatatca aaatctctgc cgcgatccct 35400
aagctcctcc ctacgaata actgtaagta ctctttcata tcctctcoga aatttttagc 35460
cataggacca ccaggaataa gattagggca agccacagta cagataaacc gaagtcctcc 35520
ccagtgcagc ttgcaaatg caagactgct ataagcatgc tggctagacc cggtgatatac 35580
ttccagataa ctggacagaa aatcacccag gcaattttta agaaaatcaa caaaagaaaa 35640
atctccagc tgacagttta gagctcggg aacaacgatg aagtaaatgc aagcggtgcg 35700
ttccagcatg gttagttagc tgatctgtaa aaaacaaaa ataaaacatt aaacctgct 35760
agcctggcga acaggtgggt aaatcgttct ctccagcacc aggcaggccca cggggtctcc 35820
ggcgcgaccc tcgtaaaaat tgcgctatg attgaaaacc atcacagaga gacgttcccg 35880
gtggccggcg tgaatgatc gacaagatga atacaccccc ggaacattgg cgtccgcgag 35940
tgaaaaaaag cgcccagga agcaataagg cactacaatg ctcagtctca agtccagcaa 36000
agcgtgcaca tgccgatgaa gcacaaaatc ctccaggtgc taaaaatgt aattactccc 36060
ctcctgcaca ggcagcgaag ccccgatcc ctccagatc acatacaaag cctcagcgtc 36120
catagcttac cgagcagcag cacacaacag gcgcaagagt cagagaaagg ctgagctcta 36180
acctgtccac ccgctctctg ctcaatatat agcccagatc tacactgacg taaaggccaa 36240
agttaaaaa taccgcca ataatcacac acgcccagca cacgcccaga aacgggtgac 36300
acactcaaaa aaatcgcgc acttcctcaa acgcccacac tgcctcatt tccgggttcc 36360
cacgtactgt catcgaatt cgactttcaa attcgtcga ccgttaaaaa cgtcaccgc 36420
cccgcccta acggtcgcgc gtctctcggc caatcacctt cctccctccc caaattcaaa 36480
cagctcattt gcatattaac gcgcacaaa agtttgaggt atattattga tgatg 36535

```

&lt;210&gt; SEQ ID NO 4

&lt;211&gt; LENGTH: 34264

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: simian adenovirus SV-1

&lt;400&gt; SEQUENCE: 4

```

tccttattct gaaacgtgc caatatgata atgagcgggg aggagcgagg cggggccggg 60
gtgacgtgcy gtgacgtggg gtgacgcggg gtggcgcgag ggcggggcgg gagtggggag 120
gcgcttagtt ttacgatg cggaaggagg ttttataccg gaagtgggt aatttgggcg 180
tatacttgta agttttgtgt aatttggcgc gaaaaccggg taatgaggaa gttgaggtta 240

```

-continued

---

|  |      |
|--|------|
| atatgtactt tttatgactg ggcggaatth ctgctgatca gcagtgaact ttgggcgctg  | 300  |
| acggggaggt ttcgctacgt ggcagtacca cgagaaggct caaaggcccc atttattgta  | 360  |
| ctcctcagcg ttttcgctgg gtatttaaac gctgtcagat catcaagagg ccaactcttga | 420  |
| gtgccggcga gtagagtttt ctccctcccg ctgcccgcat gaggctgggt cccgagatgt  | 480  |
| acgggtgttt ctgcagcgag acggccccga actcagatga gctgcttaat acagatctgc  | 540  |
| tgatgttcc caactcgcct gtggcttcgc ctccgctcgt tcatgatctt ttcgatgtgg   | 600  |
| aagtggatcc accgcaagat cccaacgagg acgcggtaaa cagtatgttc cctgaatgtc  | 660  |
| tgtttgaggc ggctgaggag ggttctcaca gcagtgaaga gagcagacgg ggagaggaac  | 720  |
| tgacttgaa atgctacgag gaatgtctgc cttctagcga ttctgaaacg gaacagacag   | 780  |
| ggggagacgg ctgtgagtcg gcaatgaaaa atgaacttgt attagactgt ccagaacatc  | 840  |
| ctggtcatgg ctgccgtgcc tgtgcttttc atagaaatgc cagcggaaat cctgagactc  | 900  |
| tatgtgctct gtgttatctg cgccttacca gcgattttgt atacagtaag taaagtgttt  | 960  |
| tcattggcgt acggtagggg attcgttgaa gtgctttgtg acttattatg tgcattatt   | 1020 |
| tctaggtgac gtgtccgacg tggaaagggga aggagataga tcaggggctg ctaattctcc | 1080 |
| ttgcactttg ggggctgtgg ttccagttgg cattttttaa ccgagtggtg gaggagaacg  | 1140 |
| agccggagga gaccgagaat ctgagagccg gcctggaccc tccagtggaa gactaggtgc  | 1200 |
| tgaggatgat cctgaagagg ggactagtgg ggggtgctagg aaaaagcaaa aaactgagcc | 1260 |
| tgaacctaga aactttttga atgagttgac tgtaagccta atgaatcggc agcgtcctga  | 1320 |
| gacgggtgtt tgactgagt tggaggatga gttcaagaag ggggaattaa acctcttga    | 1380 |
| caagtatggg tttgagcagt tgaaaactca ctggttgag ccgtgggagg atatggaaat   | 1440 |
| ggctctagac acctttgcta aagtggctct gcggccggat aaagtttaca ctattcggc   | 1500 |
| caactgttaat ataaaaaga gtgtttatgt taccggccat ggagctctgg tgcaggtgca  | 1560 |
| gacccagac cggttggtct tcaattgcgg catgcagagt ttgggccccg gggtagatag   | 1620 |
| tttgaatgga gttacatttc aaaatgtcag gtttactggt gatgatttta atggctctgt  | 1680 |
| gtttgtgact agcaccagc taaccctcca cgggttttac ttttttaact ttaacaatac   | 1740 |
| atgtgtggag tcatgggta ggggtctct gaggggctgc agttttcatg gttgctggaa    | 1800 |
| ggcggtggtg ggaagaatta aaagtgtcat gtctgtgaa aaatgcata tgaacgctg     | 1860 |
| tgtgatagct ctagcagtag aggggtacgg acggatcagg aataacgcc catctgagaa   | 1920 |
| tgatgtttt cttttgtga aaggtacggc cagcgttaag cataatatga tttgcccag     | 1980 |
| cggcctgtgc ccctcgcagc tcttaacttg cgcagatgga aactgtcaca ccttgcgcac  | 2040 |
| cgtgcacata gtgtcccact cgcgcgcac ctggccaaca tttgagcaca atatgctcat   | 2100 |
| gcgttgccgc gttcacctag gtgctagacg cggcgtgttt atgccttacc aatgtaactt  | 2160 |
| tagtcatact aagattttgc tggaaactga ttccttcct cgagtatgtt tcaatgggg    | 2220 |
| gtttgacatg tcaatggaac tttttaaagt gataagatat gatgaaacca agtctcgtt   | 2280 |
| tcgctcatgt gaatcggag ctaatcattt gaggtgtat cctgtaacc tgaacgttac     | 2340 |
| cgaggagctg aggacggacc accacatgct gtcttgcctg cgtaccgact atgaatccag  | 2400 |
| cgatgaggag tgaggtgagg ggcggagcca caaaggttat aaagggcat gagggtggg    | 2460 |
| cgcgggtgtt caaatgagc gggacgacgg acggcaatgc gttgagggg ggagtgttca    | 2520 |

---

-continued

---

gcccataatct gacatctcgt cttccttcct gggcaggagt tcgtcagaat gtagtgggct 2580  
ccaccgtgga cggacggccg gtcgcccctg caaattccgc caccctcacc tatgccaccg 2640  
tgggatcatc gttggacact gccgcggcag ctgcccgttc tgctgccgct tctactgctc 2700  
gcggcatggc ggctgatttt ggactatata accaactggc cactgcagct gtggcgcttc 2760  
ggtctctggt tcaagaagat gccctgaatg tgatcttgac tcgcctggag atcatgtcac 2820  
gtcgcctgga cgaactggct gcgcagatat cccaagctaa ccccgatacc gcttcagaat 2880  
cttaaaataa agacaacaaa atttgttgaa aagtaaaatg gctttatttg ttttttttg 2940  
ctcggtaggc tcgggtccac ctgtctcggg cgtaagac tttgtgatg ttttccaaa 3000  
cacggtacag atgggcttg atgttcaagt acatgggcat gagccatct ttggggtgga 3060  
gataggacca ctgaagagcg tcatgttccg ggggtgatt gtaaatcacc cagtcgtagc 3120  
aggggttttg agcgtggaac tggaatatgt ccttcaggag caggctaag gccaaaggta 3180  
gacccttagt gtaggtgttt acaaagcggg tgagctggga gggatgcatg cggggggaga 3240  
tgatatgcat cttggcttg attttgaggt tagctatgtt accaccagg tctctcggg 3300  
ggttcatgtt atgaaggacc accagcacgg tatagccagt gcatttgggg aacttgtcat 3360  
gcagtttga ggggaagcg tggaagaatt tagatacccc cttgtgcccc cctaggtttt 3420  
ccatgcactc atccataata atggcaatgg gaccctggc ggccgcttta gcaaacacgt 3480  
tttgggggtt gaaacatca tagttttgct cttagtgag ctcacatag gccatcttta 3540  
caaagcgggg taggaggggtg cccgactggg ggatgatagt tccatctggg cctggagcgt 3600  
agttgcctc acagatctgc atctcccagg ccttaatttc cgaggggggg atcatgtcca 3660  
cctggggggc gataaaaaac acggtttctg gcggggggtt aatgagctgg gtgaaagca 3720  
agttacgcaa cagctgggat ttgcccaac cgggtggacc gtagatgacc ccgatgacgg 3780  
gttgacgctg gtagttcaga gaggaacagc tgccgtcggg gcgcaggagg ggagctacct 3840  
cattcatcat gcttctgaca tgtttatttt cactcactaa gttttgcaag agcctctccc 3900  
caccagggga taagagtct tccaggctgt tgaagtgttt cagcggtttc aggccgtcgg 3960  
ccatgggcat cttttcaagc gactgacgaa gcaagtaag tcggtcccag agctcgggta 4020  
cgtgctctat ggaatctcga tccagcagac ttcttggttt cgggggttg gccgactttc 4080  
gctgtagggc accagccggg gggcgtccag ggccgcagg gttctgtcct tccagggtct 4140  
cagcgttcgg gtgaggggtg tctcggtgac ggtgaaggga tgagccccgg gctgggcgct 4200  
tgcgaggggt cgcttcaggc tcatctgct ggtgctgaag cgggcgtcgt ctccctgtga 4260  
gtcggccaga tagcaacgaa gcatgaggtc gtagctgagg gactcggccg cgtgtccctt 4320  
ggcgcgcagc tttcccttg aaactgtctg acatttgggt cagtgacagc acttgagggc 4380  
gtagagtttt gggccagga agaccgactc gggcgagtag gcgtcggctc cgcactgagc 4440  
gcagacggtc tcgactcca ccagccactg gagctcgggt tttagcgggat caaaaacaa 4500  
gttgctcca ttttttttga tgcgtttctt accttgctc tccatgagtc tgtgtcccgc 4560  
ttcogtgaca aaaaggtgt cggtatcccc gtagaccgac ttgagggggc gatcttccaa 4620  
aggtgttccg aggtcttccg cgtacagaa ctgggaccac tccagacaa aggctcgggt 4680  
ccaggctaac acgaaggagg cगतctgcga ggggtatctg tcgttttcaa tgagggggtc 4740  
caccttttcc aggtgtgca gacacagtc gtcctcctcc gcgtccaoga aggtgattgg 4800

---

-continued

---

cttgtaagtg taggtcacgt gacccgcacc cccccaagggt gtataaaagg gggcgtgccc 4860  
actctccccg tcactttctt ccgcatcgct gtggaccaga gccagctggt cgggtgagta 4920  
ggccctctca aaagccggca tgatttcggc gctcaagttg tcagtttcta caaacgaggt 4980  
ggatttgata ttcaagtgcc ccgcgcgcat gcttttgatg gtggaggggt ccatctgatc 5040  
agaaaacacg atctttttat tgtcaagttt ggtggcgaaa gacccgtaga gggcgttggg 5100  
aagcaacttg gcgatggagc gcagggctg atttttctcc cgatcggccc tctccttggc 5160  
ggcgatggtt agttgacgt actcgcgggc cacgcaccgc cactcgggga acacggcggg 5220  
gcgctcgtcg ggcaggatgc gcacgcgcca gcccggttg tgcaggggta tgagggtccac 5280  
gctggtggcc acctccccgc ggaggggctc gttggtcaca cacaatcgcc ccccttttct 5340  
ggagcagaac ggaggcagg gatctagcaa gttggcgggc ggggggtcgg cgtcagtggt 5400  
aaatagccg ggtagcagaa ttttattaaa ataatcgatt tcggtgtccg tgtcttgcaa 5460  
cgcgtcttcc cacttcttca ccgccagggc ccttctgtag ggattcaggg gcggtcccca 5520  
gggatggggg tgggtcaggg ccgagggcga catgcccgag atgtcgtaca cgtacagggg 5580  
ctcctcaac accccgatgt aagtggggta acagcgcgcc ccgcggatgc tggctcgcac 5640  
gtagtcgtac atctcgtgag agggagccat gagcccgtct cccaagtggg tcttgtgggg 5700  
tttttcggcc cggtagagga tctgcctgaa gatggcgtgg gagttggaag agatagtggg 5760  
gcgttgggag acgttaaagt tggtccggg cagtcccacg gagtcttggg tgaactgggc 5820  
gtaggattcc cggagcttgt ccaccagggc tgcggttacc agcacgtcga gagcgcagta 5880  
gtccaacgtc tcgcggaaca ggtttaggc cgtctcttgt ttttctccc acagttcggc 5940  
attgaggag tattcctcgc ggtctttcca gtactcttcg gcgggaaatc cttttctcgc 6000  
cgctcggtaa gaacctaaaca tgtaaaatc gttcacggct ttgtatggac aacagccttt 6060  
ttctaccggc agggcgtagc cttgagcggc ctttctgaga gaggtgtggg tgagggcgaa 6120  
ggtgtccgc accatcaact tcaggtactg atgtttgaa tccgtgtcgt cgcaggcggc 6180  
ctgttcccc acgctgaagt cgggtcgctt tttctgcctg ggattgggga gggcgaatgt 6240  
gacgtcgta aagagattt tcccggcgcg gggcatgaa ttgcgagaga tctgaaggg 6300  
tccgggcacg tccgagcgtg tgttgatgac ttgcgccgcc aggacgatct cgtogaagcc 6360  
gttgatggtg tgccccacga tgtaaagtc gataaagcgc ggctgtccct tgagggcccg 6420  
cgcttttttc aactcctcgt aggtgagaca gtccggcgag gagagacca gctccgcccg 6480  
ggcccagtcg gagagctgag ggttagccgc gagaaaagag ctccacaggt caagggctag 6540  
cagagtttgc aagcggtcgc ggaactcgcg aaactttttc cccacggcca ttttctccg 6600  
cgtcaccacg tagaaagtgc agggcggtc gttccagacg tcccatcgga gctctagggc 6660  
cagctcgcag gcttgacgaa cgagggctc ctcgcccag acgtgcatga ccagcatgaa 6720  
gggtaccaac tgtttccga acgagccat ccatgtgtag gttctacgt cgtaggtgac 6780  
aaagaccgc tgggtgcgcg cgtgggagcc gatcgggaa aagctgatct cctgccacca 6840  
gttggaggaa tgggtgttga tgtggtgaaa gtagaagtcc gcggcgca cagagcattc 6900  
gtgctgatgt ttgtaaaagc gaccgcagta gtcgcagcgc tgcacgtct gtatctctg 6960  
aatgagatgc gcttttcgcc cgcgcaccag aaaccggagg gggaggtga gacgggggct 7020  
tggtggggcg gcatcccctt cgccttggcg gtgggagtct gcgtctgcgc cctccttctc 7080



-continued

---

|            |             |             |             |             |             |      |
|------------|-------------|-------------|-------------|-------------|-------------|------|
| tgggtggacg | acggtgggga  | cgacgacgcc  | ccgggtgccg  | caagtccaga  | tctccgccac  | 7140 |
| ggagggggcg | agggcttgca  | ggaggggacg  | cagctgcccc  | ctgtccaggg  | agtcgagggc  | 7200 |
| ggccgcgctg | aggtcggcgg  | gaagcgtttg  | caagttcact  | ttcagaagac  | cgtaagagc   | 7260 |
| gtgagccagg | tgcacatggt  | acttgatttc  | caggggggtg  | ttggaagagg  | cgccaccggc  | 7320 |
| gtagaggagg | ccgtgtccgc  | gcggggccac  | caccgtgccc  | cgaggagggt  | ttatctcact  | 7380 |
| cgtcgagggc | gagcgccggg  | gggtagaggc  | ggctctgcgc  | cggggggcag  | cgagggcagt  | 7440 |
| ggcacgtttt | cgtgaggatt  | cggcagcggg  | tgatgacgag  | cccgggagact | gctggcgtgg  | 7500 |
| gcgacgacgc | ggcggttgag  | gtcctggatg  | tgccgtctct  | gcgtgaagac  | caccggcccc  | 7560 |
| cgggtcctga | acctgaaaga  | gagttccaca  | gaatcaatgt  | ctgcatcggt  | aacggcggcc  | 7620 |
| tgctgagga  | tctcctgtac  | gtcgcgccgag | ttgtcttgat  | aggcgatctc  | ggccatgaac  | 7680 |
| tgctccactt | cttcctcgcg  | gaggtcgccg  | tggcccgtct  | gctccacggg  | ggcggccagg  | 7740 |
| tcgttgagga | tgcgacgcat  | gagttgagag  | aaggcgttga  | ggccgttctc  | gttccacacg  | 7800 |
| cggctgtaca | ccacgtttcc  | gaaggagtcg  | cgcgctcgca  | tgaccacctg  | ggccacgttg  | 7860 |
| agttccacgt | ggcggcgcaa  | gacggcgtag  | tttctgaggg  | gctggaagag  | gtagttgagc  | 7920 |
| gtggtggcga | tgtgctcgca  | gacgaagaag  | tacatgatcc  | agcggccgag  | ggtcatctcg  | 7980 |
| ttgatgtctc | cgatggcttc  | gagacgctcc  | atggcctcgt  | agaagtcgac  | ggcgaagtgg  | 8040 |
| aaaaattggg | agttgcgggc  | ggccaccgtg  | agttcttctt  | gcaggaggcg  | gatgagatcg  | 8100 |
| gcgaccgtgt | cgcgcacctc  | ctgctcgaaa  | gcgccccgag  | gcgcctctgc  | ttcttcctcc  | 8160 |
| ggctcctcct | cttccagggg  | cacgggttcc  | tccggcagct  | ctgcgacggg  | gacggggcgg  | 8220 |
| cgacgtcgtc | gtctgaccgg  | cagggcgtcc  | acgaagcgtc  | cgatcatttc  | gccgcgccgg  | 8280 |
| cgacgcatgg | tctcggtgac  | ggcgcgtccg  | tttctcgagag | gtcgcagttc  | gaagaccgg   | 8340 |
| ccgcgcagag | cgcgcccgctg | cagggagggt  | aagtggttag  | ggccgtcggg  | cagggacacg  | 8400 |
| gcgctgacga | tgcattttat  | caattgctgc  | gtaggcactc  | cgtgcaggga  | tctgagaacg  | 8460 |
| tcgaggtcga | cgggatccga  | gaacttctct  | aggaaagcgt  | ctatccaatc  | gcagtcgcaa  | 8520 |
| ggtaagctga | ggacggtggg  | ccgctggggg  | gcgtcccgcg  | gcagttggga  | ggtgatgctg  | 8580 |
| ctgatgatgt | aattaaagta  | ggcggctctc  | aggcggcggg  | tggtggcgag  | gaggaccacg  | 8640 |
| tctttggggc | cggcctggtg  | aatgcgcagg  | cgctcggcca  | tgccccaggc  | ctcgcctctga | 8700 |
| cagcgacgca | ggtctttgta  | gtagtcttgc  | atcagtctct  | ccaccggaac  | ctctgcttct  | 8760 |
| ccccgtctg  | ccatgcgagt  | cgagcogaac  | ccccgcaggg  | gctgcagcaa  | cgtaggtctg  | 8820 |
| gccacgaccc | tctcggccag  | cacggcctgt  | tggatctgcg  | tgagggtggg  | ctggaagtctg | 8880 |
| tccaggtcca | cgaagcgggt  | ataggcccc   | gtgttgatgg  | tgtaggtgca  | gttggccatg  | 8940 |
| acggaccagt | tgacgacttg  | catgcggggg  | tgggtgatct  | ccgtgtactt  | gaggcgcgag  | 9000 |
| taggcgcggg | actcgaacac  | gtagtcgttg  | catgtgcgta  | ccagatactg  | gtagccaacc  | 9060 |
| aggaagtggg | gagggcgttc  | tcggtacagg  | ggccagccga  | ctgtggcggg  | ggcggcgggg  | 9120 |
| gacaggtcgt | ccagcatgag  | gcgatggtag  | tggtagatgt  | agcgggagag  | ccaggtgatg  | 9180 |
| ccggccgagg | tggtcgcggc  | cctggtgaat  | tcgcgacgc   | ggttccagat  | gttgcgcagg  | 9240 |
| ggggaaagc  | gctccatggt  | gggcacgctc  | tgccccgtga  | ggcgggcgca  | atcttgtagc  | 9300 |
| ctctagatgg | aaaaagaca   | ggcgggtcat  | cgactccctt  | ccgtagctcg  | gggggtaaag  | 9360 |

-continued

---

|             |            |             |            |            |             |       |
|-------------|------------|-------------|------------|------------|-------------|-------|
| tcgcaagggt  | gcggcggcgg | ggaaccccg   | ttcgagaccg | gccggatccg | ccgctcccga  | 9420  |
| tgcgcctggc  | cccgcaccca | cgacgtccgc  | gtcgagaccc | agccgcgacg | ctccgccccca | 9480  |
| atacggaggg  | gagtcttttg | gtgttttttc  | gtagatgcat | ccggtgctgc | ggcagatgcg  | 9540  |
| acctcagacg  | cccaccacca | ccgcccgccg  | ggcagtaaac | ctgagcggag | gcggtgacag  | 9600  |
| ggaggaggag  | gagctggcct | tagacctgga  | agagggagag | gggctggccc | ggctggggagc | 9660  |
| gccgtcccca  | gagagacacc | ctagggttca  | gctcgtgagg | gacgccaggg | aggcttttgt  | 9720  |
| gccgaagcag  | aacctgttta | gggaccgcag  | cggtcaggag | gcggaggaga | tgcgcgattg  | 9780  |
| caggtttcgg  | gcgggtagag | agctgagggc  | ggccttcgat | cgggagcggc | tcctgagggc  | 9840  |
| ggaggatttc  | gagcccagc  | agcgttctgg  | ggtgagccc  | gcccgcgctc | acgtctcggc  | 9900  |
| ggccaacctg  | gtgagcgcgt | acgagcagac  | ggtgaacgag | gagcgcaact | tccaaaagag  | 9960  |
| ctttaacaat  | cacgtgagga | ccctgatcgc  | gagggaggag | gtgaccatcg | ggctgatgca  | 10020 |
| tctgtgggac  | ttcgtggagg | cctacgtgca  | gaacccggcc | agcaaacctc | tgacggccca  | 10080 |
| gctgttcctg  | atcgtgcagc | acagccgcga  | caacgagacg | ttccgcgacg | ccatgttgaa  | 10140 |
| catcgccggag | cccgagggtc | gctggctctt  | ggatctgatt | aacatcctgc | agagcatcgt  | 10200 |
| ggtgcaggag  | aggggctca  | gcttagcggg  | caagggtggc | gccattaact | attcogatgca | 10260 |
| gagcctgggg  | aagttctacg | ctcgcgaagat | ctacaagagc | ccttacgtgc | ccatagacaa  | 10320 |
| ggaggtgaag  | atagacagct | tttacatcgc  | catggcgctg | aaggtgctga | cgctgagcga  | 10380 |
| cgacctcggc  | gtgtaccgta | acgacaagat  | ccacaaggcg | gtgagcgcga | gccgccggcg  | 10440 |
| ggagctgagc  | gacagggagc | tgatgcacag  | cctgcagagg | gcgctggcgg | gcgccgggga  | 10500 |
| cgaggagcgc  | gaggcttact | tcgacatggg  | agccgatctg | cagtggcgtc | ccagcgcgcg  | 10560 |
| cgcttgagg   | gcggcgggct | accccgacga  | ggaggatcgg | gacgatttgg | aggaggcagg  | 10620 |
| cgagtacgag  | gacgaagcct | gaccgggcag  | gtgttgtttt | agatgcagcg | gccggcggac  | 10680 |
| ggggccaccg  | cggatcccgc | acttttgcca  | tccatgcaga | gtcaaccttc | ggcgtgacc   | 10740 |
| gcctccgatg  | actggcggcg | ggccatggac  | cgcattatgg | cgctgactac | ccgcaacccc  | 10800 |
| gaggctttta  | gacagcaacc | ccaggccaac  | cgtttttcgg | ccatcttggg | agcggtggtg  | 10860 |
| ccctcccgca  | ccaaccccac | acacgagaaa  | gtcctgacta | tcgtgaaocg | cctggtagac  | 10920 |
| agcaaggcca  | tcgcccgcga | cgaggcgggc  | ttgatttaca | acgctctgct | ggaacgggtg  | 10980 |
| gcgcgtaca   | acagcactaa | cgttcagacc  | aatctggatc | gcctcaccac | cgacgtgaag  | 11040 |
| gaggcgctgg  | ctcagaagga | gcggtttctg  | agggacagca | atctgggctc | tctggtggca  | 11100 |
| ctcaacgcct  | tcctgagcac | gcagccggcc  | aacgtgcccc | gcgggcagga | ggactacgtg  | 11160 |
| agcttcatca  | gcgctctgag | gctgctggtg  | tccgaggtgc | cccagagcga | ggtgtatcag  | 11220 |
| tctgggcccg  | attacttctt | ccagaogtcc  | cgacagggct | tgcaaacggt | gaaactgact  | 11280 |
| caggccttta  | aaaacttgca | aggcatgtgg  | ggcgttaagg | ccccggtggg | cgatcgagcc  | 11340 |
| accatctcca  | gtctgctgac | ccccaacact  | cgcctgctgc | tgctcttgat | cgcgccgttc  | 11400 |
| accaacagta  | gcactatcag | ccgtgactcg  | tacctgggtc | atctcatcac | tttgtaccgc  | 11460 |
| gaggccatcg  | gtcaggctca | gatcgacgag  | cacacatata | aggagatcac | taacgtgagc  | 11520 |
| cgggcctcgg  | gtcaggaaga | taccggcagc  | ctggaagcca | cgttgaactt | tttgcatacc  | 11580 |
| aaccggaggc  | aaaaaatacc | ctcccagttt  | acgttaagcg | ccgaggagga | gaggattctg  | 11640 |

---

-continued

---

cgatacgtgc agcagtcctg gagtctgtac ttgatgcggg agggcgccac cgcttccacg 11700  
gcttttagaca tgacggctcg gaacatggaa ccgtcctttt actccgcccc cgggccgttc 11760  
attaaccgtc tgatggacta cttccatcgc gcggccgcca tgaacgggga gtacttcacc 11820  
aatgccatcc tgaatccgca ttggatgccc cgtccggct tctacaccgg cgagtttgac 11880  
ctgcccgaag ccgacgacgg ctttctttgg gacgacgtgt ccgacagcat tttcacgcgg 11940  
ggcaatcgcc gattccagaa gaaggagggc ggagacgagc tccccctctc cagcgtggag 12000  
gcggcctcta ggggagagag tccctttccc agtctgtctt ccgccagcag tggtcgggta 12060  
acgcgcccg ggttgccggg ggagagcgc tacctgaacg accccttget gcggccggct 12120  
aggaagaaaa atttcccaa caacggggtg gaaagccttg tggataaaat gaatcgttg 12180  
aagacctacg ccagagagca gcgggagtg gaggacgtc agccgcgacc gctggttccg 12240  
ccgcactggc gtcgtcagag agaagaccgg gacgactccg cagacgatag tagcgtgttg 12300  
gacctgggag ggagcggagc caacccttt gctcacttgc aaccaaggg gcgttccagt 12360  
cgctctact aataaaaaag acgcggaaac ttaccagagc catggccaca gcgtgtgtcc 12420  
tttcttctc tctttcttc tcggcgggc agaatgagaa gagcgtgag agtcacgcgg 12480  
gcggcgtatg agggccgcc cccttcttac gaaagcgtga tgggatcagc gaacgtgccg 12540  
gccacgttg aggcgcctta cgttctctcc agatacctgg gacctacgga gggcagaaac 12600  
agcatccgtt actccgagct ggcacccctg tacgatacca ccaaggtgta cctggtggac 12660  
aacaagtcgg cggacatcgc ctccctgaat tatcaaaacg atcacagcaa tttctgact 12720  
accgtggtgc agaacaatga cttcaccctg acggaggcgg gcacgcagac cattaacttt 12780  
gacgagcgtt cccgctgggg cggtcagctg aaaaccatcc tgcacaccaa catgcccaac 12840  
atcaacgagt tcatgtccac caacaagttc agggccaggc tgatggttaa aaaggctgaa 12900  
aaccagcctc ccgagtacga atggtttgag ttcaccattc ccgagggcaa ctattccgag 12960  
accatgacta tcgatctgat gaacaatcgc atcgtggaca attacctgca agtggggagg 13020  
cagaacgggg tatttgaaag cgatatcggc gtaaaatttg ataccagaaa cttccgactg 13080  
gggtgggatc ccgtgaccaa gctggtgatg ccaggcgtgt acaccaacga ggcttttac 13140  
cccgacatcg tgctgtgccc ggggtgcggg gtggacttca ctcagagccg tttgagtaac 13200  
ctgttaggga tcagaaagcg ccgccccttc caagaggct ttcagatcat gtatgaggac 13260  
ctggaaggag gtaacattcc aggtttgcta gacgtgccgg cgtatgaaga gagtgttaa 13320  
caggcggag cgcaggagc agagattcga ggcgacacct ttgccacgga acctcacgaa 13380  
ctggaataa aacctctgga acaagacagt aaaaaacgga gttacaacat tatatccggc 13440  
actatgaata cttgtaccg gagctggttt ctggcttaca actacgggga tcccgaag 13500  
ggagtgagat catggacct actcaaccac acggacgtga cctgcggctc gcagcaagt 13560  
tactgtccc tgccgatat gatgcaagac ccggtcacct tccgcccctc cacccaagtc 13620  
agcaacttcc cgggtggtgg caccgagctg ctgcccgtcc atgccaagag cttctacaac 13680  
gaacaggccg tctactcga actcattcgc cagtccaccg cgcttaccga cgtgttcaat 13740  
cgctttcccg agaaccagat tctggtgcgc cctcccgtc ctaccattac caccgtcagt 13800  
gaaaacgttc ccgccctcac agatcacgga accctgccgc tgcgcagcag tatcagtgga 13860  
gttcagcgcg tgaccatcac cgacgccaga cgtcgaacct gtcctactgt ttacaaagct 13920

---

-continued

---

cttggcgtag tggctcctaa agtgetctct agtcgcacct tctaacaatg tccatcctca 13980  
tctctccccg taacaacacc ggctggggac tgggctccgg caagatgtac ggcggagcca 14040  
aaaggcgctc cagtcagcac ccagttcgag ttcggggcca cttccgtgct ccctggggag 14100  
cttacaagcg aggactctcg gcccgaacgg cggtagacga taccatagat gccgtgattg 14160  
ccgacgcccc ccggtacaac cccggaccgg tcgctagcgc cgcctccacc gtggattccg 14220  
tgatcgacag cgtggtagct ggcgctcggg cctatgctcg ccgcaagagg cgctgcac 14280  
ggagacgtcg ccccaccgcc gccatgctgg cagccagggc cgtgctgagg cgggcccgga 14340  
gggtaggcag aagggtatg cgcgcgctg cgcacaacgc cgcgcggggg agggcccgcc 14400  
gacaggctgc ccgccaggct gctgcgcca tcgctagcat gcccagacc aggagaggga 14460  
acgtgtactg ggtgcgcat tctgtgacgg gaggccgagt gccggtgcgc agccgacctc 14520  
ccccgaagta gaagatcaa gctgcgaaga cggcggtact gagtctcctt gttgttatca 14580  
gccccaatg agcaagcgca agtttaaga agaactgctg cagacgctgg tcctgagat 14640  
ctatggccct ccggacgtga agcctgacat taagccccgc gatatacagc gtgttaaaaa 14700  
gcgggaaaaa aaagaggaac tcgcggtggt agacgatggc ggagtggaat ttattaggag 14760  
tttcgccccg cgacgcaggg ttcaatgaa agggcgggcg gtacaacgcg ttttgaggcc 14820  
gggcaccgcg gtagttttta cccgggaga gcggtcggcc gttaggggtt tcaaaaggca 14880  
gtacgacgag gtgtacggcg acgaggacat attggaacag gcgctcaac agatcggaga 14940  
atgtgcctac gaaagcggtt cgcgctcga agacctggcc atcgctttag acagcgcaa 15000  
ccccacgcc agcctcaaac ctgtgacgct gcagcaggtg ctccccgtga gcgccagcac 15060  
ggacagcaag aggggaataa aaagagaaat ggaagatctg cagcccacca tccagctcat 15120  
ggtccctaaa cggcagaggc tggaagaggc cctggagaaa atgaaagtgg acccaagcat 15180  
agagccggac gtcaaatgca gccgatcaa agaagtggcc cctggtctcg gggtcagac 15240  
gggtgatata cagatccccg tcacgtcagc ttcgaccgcc gtggaagcca tggaaacgca 15300  
aacgaaacc cctgccgcga tcggtaccag ggaagtggcg ttgcaaaccg acccctggta 15360  
cgaatacgcc gccctcggc gtcagaggcg acccgctcgt tacggccccg ccaacgcat 15420  
catgccagaa tatgcgctgc atccgtctat cctgccacc cccggtacc ggggagtgc 15480  
gtatgccccg tcaggaaccc gccgcgaac ccgctcggc cgcgctccc gtcgtgctct 15540  
ggcccccggt tcggtcggcc gcgtaacacg ccggggaaag acagttacca ttccaaccc 15600  
gcgctaccac cctagcatcc tttaatgact ctgcccgttt gcagatggct ctgactgcc 15660  
gcgtgcgctt tcccgttccg cactatcgag gaagatctcg tcgtaggaga ggcatggcg 15720  
gtagtggctg ccggcgggct ttgctcagc gcataaaagg cggaaattta cccgctctga 15780  
taccataat cgcgcggcc atcggtgcca taccggcgct cgttcagtgc gccttgcaag 15840  
cagctcgtaa taaataaacg aaggcttttg cacttatgct ctggtcctga ctattttatg 15900  
cagaaagagc atggaagaca tcaattttac gtcgctggct ccgcggcacg gctcgggcc 15960  
gctcatgggc acctggaac acatcggcac cagtcagctc aacggggggc ctttcaattg 16020  
ggggagcctt tggagcgcca ttaaaaactt tggctccacg attaaatcct acggcagcaa 16080  
agcctggaac agtagtctg gtcagatgct ccgagataaa ctgaaagaca ccaacttcca 16140  
agaaaaagtg gtcaatgggg tggtagccgg catccacggc gcggtagatc tcgccaacca 16200

---

-continued

---

agcggtgacg aaagagattg acaggcgctt ggaaagctcg cgggtgccgc cgcagagagg 16260  
ggatgaggtg gaggtcgagg aagtagaagt agaggaaaag ctgccccgcg tggagaaagt 16320  
tcccgggtgcg cctccgagac cgcagaagcg acccaggcca gaactagaag aaactctggt 16380  
gacggagagc aaggagcctc cctcgtacga gcaagccttg aaagagggcg cctctccacc 16440  
ctaccaaatg acaaaaccga tcgcgctat ggctcggccg gtgtacggga aggactacaa 16500  
gcctgtcacg ctagagctcc ccccgcgccc accgcccgc cccacgcgccc cgaccgttcc 16560  
cccccccctg cgggtccctg cggcgggacc cgtgtccgca cccgtcgcg tgcctctgcc 16620  
agccgcccg ccaagtggcg tggccactgc cagaaacccc agaggccaga gaggagccaa 16680  
ctggcaaagc acgctgaaca gcatcgtggg cctgggagtg aaaagcctga aacgccgccc 16740  
ttgctattat taaaagtga gctaaaaat ttcccgttgt atacgcctcc tatggtaccg 16800  
ccagagagcg gtgactgtcg ccgagagcgc cgctttcaag atggccaccc catcgatgat 16860  
gccgcagtgg tcttacatgc acatcgccgg gcaggacgcc tcggagtacc tgagccccgg 16920  
tctcgtgcag ttcgcccgcg ccaccgacac ctacttcagc ttgggaaaca agtttagaaa 16980  
ccccaccgtg gccccaccc acgatgtaac cacggaccgc tcgcaaaggc tgaccctgcg 17040  
ttttgtgccc gtagaccggg aggacaccgc gtactcttac aaagtgcgct acacgctggc 17100  
cgtaggggac aaccgagtgc tggacatgac cagcacctac tttgacatcc ggggagtget 17160  
ggatcgcggt cccagtttta agccctactc ggttaccgcg tacaattccc tggctcccaa 17220  
ggcgctccc aaccctgcag aatggacgaa ttcagacagc aaagttaaag tgagggcaca 17280  
ggcgcccttt gttagctcgt atggtgctac agcgattaca aaagagggta ttcagggtgg 17340  
agtaacctta acagactccg gatcaacacc acagtatgca gataaacgt atcagcctga 17400  
gccgcaaatt ggagaactac agtgaacag cgatgttggg accgatgaca aaatagcag 17460  
aagagtgcata aagaaaaaa cgcctatggt cccttggtac ggctcatatg ccaggccac 17520  
taatgaaaaa ggaggacag caacaccgct cgctagtcaa gacgtgcaa atcccgaatt 17580  
acaatttttt gcctctacta atgtcgccaa tacaccaaaa gcagttctat atgcgagga 17640  
cgtgtcaatt gaagcgcag acaactcact ggtgttcaa ccaacagtca ctgaagcat 17700  
tacaagtta gaggctctac tgaccacaac agctgctccc aaccgtcaa actacatagc 17760  
ctttagagat aattttattg gtctcatgta ctacaatagc acaggtaaca tgggagtact 17820  
ggcaggccag gcttctcagc taaatgcagt tgttgacctg caagacagaa atactgagct 17880  
gtctaccaa ctcatgttg acgcccctcg agaccgcagt cggctcttt ctatgtgaa 17940  
ccaagctgtg gatagttacg atcctgatgt aagaatcata gaaaacctg gcgtagaaga 18000  
tgaattgcct aattattgct ttcctttggg aggcatggca gtaaccgaca cctactgccc 18060  
tataaagggt aatggaggag gcaatggatg ggaagccaat aacggcgctt tcaccgaaag 18120  
aggagtggaa ataggttcag ggaacatggt tgccatggag attaacctgc aagccaacct 18180  
atggcgtagc tttctgtact ccaatattgg gctgtacctg ccagactctc tcaaaatcac 18240  
tcctgacaac atcacactcc cagagaacaa aaacacctat cagtatatga acggtcgcgt 18300  
gacgccaccc gggctggttg acacctactg taacgtgggc gcgctggtt ccccgatgt 18360  
catggacagt attaacctt ttaatoacca ccgcaacgcc ggactccgct accgttccat 18420  
gctcctggga aacggagcgt acgtgocctt ccacatccag gtgcccaga aattctttgc 18480

---

-continued

---

aattaaaaac ctgctgctgc tccccggttc ctacacctac gagtggaact tccgcaagga 18540  
cgtgaacatg atctttgcaga gctcgtctgg caatgacctg cgagtggacg gggccagcat 18600  
ccgcttcgac agcatcaacc tgtacgcaa ctttttcccc atggcccaca acacggcctc 18660  
caccctggaa gccatgctgc gcaacgacac caacgaccaa tctttcaacg actacctgtg 18720  
cgcggccaac atgctgtacc ccattccccg caacgcccacc agcgtgccc a tctccattcc 18780  
ctctcgcaac tgggcagcct tcaggggctg gagtttcacc cgcctcaaaa ccaaggagac 18840  
ccccctgctg ggctccgggt tcgacccccta cttcgtctac tccggctcca tcccctacct 18900  
ggacggcacc ttctacctca accatacttt caaaaagtg tcaatcatgt tcgactcctc 18960  
cgtcagctgg cccggcaacg accgtctgct gacgcccac gagttcgaaa tcaagcgttc 19020  
ggtggacggt gaagggtaca acgtggctca gagcaacatg accaaggact ggttcctgat 19080  
tcagatgctc agccactaca acatcggtca ccagggttc tacgtgcccg aaaattacaa 19140  
ggaccgcatg tactctttct tcagaaaactt ccaaccatg agccgcaaaa ttgtagattc 19200  
aacggcttac actaattatc aggatgtgaa actgccatac cagcataaca actcagggtt 19260  
cgtggctac atgggaccca ccattcgaga ggggcaggcc taccggcca actatcccta 19320  
tcccctgatt gggccaccg ccgtgcccag cctcacgacg aaaaagtcc tctgacccg 19380  
ggtgatgtgg aggatcccct tctctagcaa cttcatgtct atgggctccc tcaccgacct 19440  
ggggcagaac atgctgtacg ccaactccgc tcacgcttg gatatgacct ttgaggtgga 19500  
tcccattgat gagcccacgc ttctctatgt tctgtttgaa gtcttcgacg tggtgccat 19560  
ccaccagccg caccgcgccg tcattcgaggc cgtctacctg cgcacacctt tctctgcccg 19620  
taaccgccc acctaagaa gccgatgggc tccagcgaac aggagctgca ggccattgtt 19680  
cgcgacctg gctgcccggc ctactttttg ggcacctcg acaagcgttt tcccggett 19740  
atgtccccc acaagccgct ctgtgccatc gttaacacgg ccggacggga gaccggggg 19800  
gtccactggc tcgctctcgc ctggaacccg cgtaaccgca cctgtaacct gttogacct 19860  
tttggtttct ccgacgaaag gctgaagcag atctaccagt tcgagtacga ggggctcctc 19920  
aagcgcagcg ctctggcctc caccgcccac cactgctca ccctggaaaa gtccacccaa 19980  
acggtccag gggccctctc gggcctctc gggctctct gttgatgtt tttgacgcc 20040  
ttcgtgcact ggctccacac cccatgat cacaacccca ccatggatct gctcaccgga 20100  
gtgccaaca gcatgctta cagccccag gtcgcccga ccctgcccg taaccaggaa 20160  
cacctgtatc gctttctgg gaaactctt gctattttc gccgccaccg gcagcgcac 20220  
gaacgggcca cggccttcga aagcatgagc caaagagtgt aatcaataaa aaacattttt 20280  
atgtgacatg atacgcgctt ctggcgtttt attaaaaatc gaagggttcg agggaggggt 20340  
cctgctgcc gctggggagg gacacgttg gatactgaa acgggcgctc caacgaaact 20400  
cggggatcac cagccgcccg aggggcact cttctagggt ctgcttcaa aactgcccga 20460  
ccagctgcag ggctccatg acgtcggcg ccgatattt gaagtgcag ttagggcccg 20520  
agctcccgc gctgttcgg aacacgggtt tggcacactg gaacaccagc acgcccgggt 20580  
tgtggatact ggccagggcc gtcgggtcgg tcacctccga cgcattcaga tcctcggcgt 20640  
tgctcagggc aaacggggct agcttgaca tctgcccgc aatctgggtt actaggtcgc 20700  
gcttgttgag gcagtcgag cgcagaggga tcaggatgc tcgctgcccg cgttgcatga 20760

---

-continued

---

taggtaact cgccgccagg aactcctcca ttgacggaa ggccatctgg gctttgccgc 20820  
cctcgggtga gaatagcccg caggacttgc tagagaatac gttatgaccg cagttgacgt 20880  
cctccgcgca gcagcgggcy tcttcgttct tcagctgaac cacgttgcgg ccccaacggt 20940  
tctggaccac cttggctcta gtgggggtgct ccttcagcgc ccgctgtccg ttctcgctgg 21000  
ttacatccat ttccaacacg tgctccttgc agaccatctc cactccgtgg aagcaaaaca 21060  
ggacgcctc ctgctgggta ctgcatgctc ccatacggc gcattccggtg ggctcccagc 21120  
tcttgtgttt tacccccgcg taggettcca tgtaagccat aaggaatctg cccatcagct 21180  
cggtaagggt cttctggttg gtgaaggta gcggcaggcc gcggtgctcc tcgttcaacc 21240  
aagtttgaca gatcttgcgg tacacgcctc cctggtcggg cagaaactta aaagccgctc 21300  
tgctgtcgtt gtctacgtgg aacttctcca ttaacatcat catggtttcc atacccttct 21360  
cccacgctgt caccagtggt ttgctgtcgg ggttcttcac caacacggcg gtagaggggc 21420  
cctcgcggc cccgacgtcc ttcattggtca ttctttgaaa ctccacggag ccgtccgcgc 21480  
gacgtactct gcgcaccgga gggtagctga agcccacctc caccacggtg ccttcgccct 21540  
cgctgtcggg gacaatctcc ggggatggcg gcggcggggg tgtcgccttg cgagccttct 21600  
tcttgggagg gagctgaggc gcctcctgct cgcgctcggg gctcatctcc cgcaagtagg 21660  
gggtaatgga gctgcctgct tggttctgac ggttggccat tgtatcctag gcagaaagac 21720  
atggagctta tgcgcgagga aactttaacc gcccctccc ccgtcagcga cgaagatgtc 21780  
atcgtcgaac aggaccggg ctacgttacg ccgcccagg atctggaggg gcctgaccgg 21840  
cgcgacgcta gtgagcggca ggaatgag aaagaggagg cctgctacct cctggaaggc 21900  
gacgttttgc taaagcattt cgccaggcag agcaccatag ttaaggaggc cttgcaagac 21960  
cgctccgagg tgcccttgga cgtcgcggcg ctctcccagg cctacgaggc gaacctttc 22020  
tcgctcggag tgccctcga gagacagccc aacggcacct gcgagccca cccgcgactc 22080  
aacttctacc ccgtgttcgc cgtaccagag gcgctggcca cctatcacat ttttttcaa 22140  
aaccaacgca tccccctatc gtgccgggccc aaccgcaccg cggccgatag gaatctcagg 22200  
cttaaaaacg gagccaacat acctgatatc acgtcgtctg aggaagtgc caagattttc 22260  
gaggtctctg gtcgagatga gaagcgggcy gcgaacgctc tgcagaaaga acagaaagag 22320  
agtcagaacg tgctggtgga gctggagggg gacaacgcgc gtctggccgt cctcaaacgc 22380  
tgcatagaag tctcccactt cgcctacccc gccctcaact tgcacccaa agttatgaaa 22440  
tcggtcatgg atcagctgct catcaagaga gctgagcccc tggatccoga ccaccccgag 22500  
gcggaact cagaggacgy aaagcccgtc gtcagcagcy aggagctcga gcggtggctg 22560  
gaaaccaggg accccaaca gttgcaagag aggcgcaaga tgatgatggc ggcctgctg 22620  
gtcaccgtgg agctggaatg cctgcaacgy tttttcagcy acgtggagac gctacgcaa 22680  
atcggggaat ccctgacta cacctccgc cagggtctag tccgccaggc ctgcaagatc 22740  
tccaacgtgg agctcagcaa cctggtctcc tacatgggca tcctccacga gaaccggctg 22800  
ggcagagcy tgctgactg caccttcaa ggcgaggcgc ggcgggacta cgtgagagac 22860  
tgcatctacc tcttctcacc cctcacctgg cagaccgcca tggcgtctg gcagcagtgc 22920  
ttggaagaga gaaacctcaa agagctagac aaactcctct gccgccagcy gcgcgcctg 22980  
tggtccggtt tcagcagcy cacggtcgc agcgtctgcy cggacatcat ctcccggag 23040

---

-continued

---

cgctgatga aaaccttgca aaacggcctg ccggatttca tcagtcaaag ctttttgcaa 23100  
aacttccgct cttttgtcct ggaacgctcc gggatcttgc ccgccatgag ctgcgcgcta 23160  
ccttctgact ttgtcccccct ctcctaccgc gagtgccctc ccccaactgtg gagccactgc 23220  
tacctcttcc aactggccaa ctttctggcc taccactccg acctcatgga agacgtaagc 23280  
ggagagggtt tactggagtg ccaactgccgc tgcaacctgt gcacccccca cagatcgctg 23340  
gcctgcaaca ccgagctact cagcgaaacc caggtcatag gtaccttoga gatccagggg 23400  
ccccagcagc aagaggggtg ttccggcttg aagctcactc cggcgtgtg gacctcggt 23460  
tacttacgca aattttagc cgaggactac cacgcccaca aaattcagtt ttacgaagac 23520  
caatctcgac caccgaaagc cccctcagc gcctgctca taccagag caagatcctg 23580  
gccccattgc aatccatcaa ccaagcgcgc cgcgatttcc tttgaaaaa gggtcgggg 23640  
gtgtacctg accccagac cggcgaggaa ctcaaccctg ccacactctc cgtcgaagca 23700  
gccccccga gacatgccgc ccaagggaa cgcgaagcag ctgatcgtc gccagagagc 23760  
gaagaagcaa gagctctcc agcagcaggt ggaggacgag gaagagatgt gggacagcca 23820  
ggcagaggag gtgtcagagg acgaggagga gatggaaagc tgggacagcc tagacagga 23880  
ggaggacgag ctttcagagg aagagggcag cgaagaaaa ccacctgcat ccagcgcgc 23940  
ttctctgagc cgacagccga agccccggcc ccgacgccc ccggccggct cactcaaagc 24000  
cagccgtag tgggacgcca ccgaatctcc agcggcagcg gcaacggcag cgggtaaggc 24060  
caaacgcgag cggcgggggt attgtcctg cggggccac aaaagcagta ttgtgaactg 24120  
cttgcaaac tgcgggggaa acatctcctt tgcccagcgc tacctcctt tccatcacg 24180  
tgtggccttc cctcgcaacg ttctctatta ttaccgtcat ctctacagcc cctacgaaac 24240  
gctcggagaa aaaagctaag gcctcctccg ccgcgaggaa aaactccgcc gccgctgccg 24300  
ccgcaagga tccaccggcc accgaagagc tgagaaagcg catcttccc actctgtatg 24360  
ctatctttca gaaagccgc gggcagcacc ctacgcgcga actgaaaata aaaaaccgct 24420  
ccttccgctc gctcaccgc agctgtctgt accacaagag agaagaccag ctgcagcgca 24480  
ccctggacga cgcggaagca ctgttcagca aatactgctc agcgtctctt aaagactaaa 24540  
agaccgcgc tttttcccc tcggcgcca aaaccacgt catcgcagc atgagcaagg 24600  
agattccac cccctacatg tggagctatc agcccagat gggcctggcc gcgggggccc 24660  
cccaggacta ctccagcaag atgaactgac tcagcgcgg cccccacatg atctcacgag 24720  
ttaacggcat ccgagccca cgaaaaccaga ttctcttaga acaggggca atcaccgcca 24780  
caccgccgc ccaactcaac ccgcctagtt ggcccgcgc ccagggtgat caggaaaatc 24840  
cccgccgac cacagtccct ctgccacgc acgcgaggc cgaagtctc atgactaact 24900  
ctgggtaca attagcggc gggctcaggt acgccagta cagaggtcgg gccgtcctt 24960  
actctcccg gagtataaag agggatgatc ttcgaggcc aggtatccag ctcaacgacg 25020  
agacggtgag ctctcaacc ggtctcagac ctgacggagt cttccagctc ggaggagcgg 25080  
gccgtcttc cttcaccact cgcaggcct acctgacct gcagagctct tctcgcagc 25140  
cgcgtccgg gggaatcgc actctccagt tcgtggaaga gttcgttccc tccgtctact 25200  
tcaaccctt ctccggctc cctggacgt acccgagcgc cttcattccc aactttgacg 25260  
cagtgagtga atccgtgac ggctaogact gatgacagat ggtcgggccc tgagagctcg 25320



---

-continued

---

gctgcgacat ctgcatcact gccgtcagcc tcgctgctac gctcgggagg cgatcgtctt 25380  
cagctacttt gagctgccg acgagcacc ctaggggtccg gctcacgggt tgaactcga 25440  
gatcgagaac gcgctcgagt ctcgcctcat cgacaccttc accgcccgac ctctcctggt 25500  
agaaatccaa cgggggatca ctaccatcac cctgttctgc atctgcccga cgcccggatt 25560  
acatgaagat ctgtgtgttc atctttgcgc tcagttaaat aaaaactgaa ctttttgccg 25620  
caccttcaac gccatctgtg atttctacaa caaaaagttc ttctggcaaa ggtacacaaa 25680  
ctgtatttta ttctaattct acctcatcta tcgtgctgaa ctgcgctgc actaacgaac 25740  
ttatccagtg gattgcaaac ggtagtgtg gcaagtactt ttgggggaac gatatagtta 25800  
gtagaaataa cagcctttgc gagcactgca actcctccac actaatcctt tatccccat 25860  
ttgttactgg atggatatag tgcgttggt cgggtttaaa tcctagtgc ttccataagt 25920  
ggtttctaca aaaagagacc cttcccaca attctgttc tttttcgc ctatcctact 25980  
gctgttctcc ctctggttac tctttcaaac ctctaattgg tattttagct ttgatactca 26040  
taatctttat taactttata ataattaaca acttacagta aacatgcttg ttctactgct 26100  
cgccacatct ttcgctctct ctcacgccag aacaagtatt gttggcgag gttacaatgc 26160  
aactctcaa tctgcttaca tgccagatc cgaccagata ccccatatta cgtggtactt 26220  
acaaacctcc aaacctaat cttcatttta tgaaggaaac aaactctgcg atgactccga 26280  
caacagaacg cacacatttc cccacccttc actacaattc gaatgcgtaa acaaaaagctt 26340  
gaagctttac aacttaaagc cttcagatc tggctgttac catgctgtag ttgaaaaaag 26400  
taatttagaa gtccacagtg attacattga attgacggt gtggacctgc cacctccaaa 26460  
atgtgaggtt tcctcctctt acctgaagt tcaagcgtg gatgcctact gcctcataca 26520  
cattaactgc agcaactcta aatatccagc tagaatttac tataatggac aggaaagtaa 26580  
tcttttttat ttttaacaa caagcgtgg taacggtaaa cagttacctg actattttac 26640  
tgctgtgtgt gaattttcca cctacagaga aacgtatgcc aagcggcctt acaatttctc 26700  
ataccggtt aacgacctt gcaatgaaat acaagcgtc gaaactgaa ctgattttac 26760  
tccaattttc attgctgcca ttgttgaag cttaattacc attattgtca gcctagcatt 26820  
ttactgcttt tacaagccca aaaacctaa gtttgaaaa cttaactaa aacctgtcat 26880  
tcaacaagtg tgattttgtt ttccagcatg gtagctgcat ttctacttct cctctgtcta 26940  
cccatcattt tcgtctcttc aactttcgc gcagtttccc acctggaacc agagtgccta 27000  
ccgcttttg acgtgtatct gattctcacc tttgtttgtt gtatatcoat ttgcagtata 27060  
gcctgctttt ttataacaat ctttcaagcc gccgactatt tttacgtgcg aattgcttac 27120  
ttagacacc atcctgaata cagaaatcaa aacgttgct ccttactttg ttggcatga 27180  
ttaagttatt gctgatactt aattatttac cctaatacaa ctgtaattgt ccattcacca 27240  
aacctggtc attctacacc tgttatgata aaatccccga cactcctggt gcttgcttt 27300  
acgcagccac cgccgcttg gtatttatat ctacttgct tggagtaaaa ttgtatttta 27360  
ttttacacac tgggtggcta catcccagag aagatttacc tagatacct cttgtaaacg 27420  
cttttcaatt acagcctctg cctcctcctg atcttcttcc tcgagctccc tctattgtga 27480  
gctactttca actcaccggt ggagatgact gactctcagg acattaatat tagtgtggaa 27540  
agaatagctg ctacagctca gcgagaaacg cgagtgttg aatacctgga actacagcaa 27600

---

-continued

---

cttaaagagt cccactggtg tgagaaagga gtgctgtgcc atgttaagca ggcagccctt 27660  
tcctacgatg tcagcgttca gggacatgaa ctgtcttaca ctttgccttt gcagaaacaa 27720  
accttctgca coatgatggg ctctacctcc atcacaaatca cccaacaagc cgggcctgta 27780  
gagggggcta tcctctgtca ctgtcacgca cctgattgca tgtccaaact aatcaaaact 27840  
ctctgtgctt taggtgatat ttttaagggtg taaatcaata ataaacttac cttaaatttg 27900  
acaacaaatt tctggtgaca tcattcagca gcaccacttt accctcttcc cagctctcgt 27960  
atgggatgag atagtggtg gcaaacttcc tccaaacctt aaaagaaata ttgggatcca 28020  
cttctctgtc ctccccaca attttcatct tttcatagat gaaaagaacc agagttgatg 28080  
aagactcaa ccccgtctac ccctatgaca ccacaaccac tcctgcagtt ccctttatat 28140  
cacccccctt tgtaaacagc gatggcttc aggaaaacc cccaggtggt ttaagtctgc 28200  
gaatagctaa acccctatat ttcgacatgg agagaaaact agccctttca cttggaagag 28260  
ggttgacaat taccgcccgg ggacaattag aaagtacgca gagcgtacaa accaaccac 28320  
cgttgataat taccaacaac aacacactga ccctacgca tcttcccccc ttaaacctaa 28380  
ctgacaatag cttagtgtca ggctactcga gtcctctccg cgtcacagac aacaaactta 28440  
catttaactt cacatcacca ctccggtatg aaaatgaaaa cttactttt aactatacag 28500  
agcctcttaa acttataaat aacagccttg ccattgacat caattcctca aaaggcctta 28560  
gtagcgtcgg aggctcacta gctgtaaacc tgagttcaga cttaaagttt gacagcaacg 28620  
gatccatagc ttttggcata caaacctgt ggaccgctcc gacctcgact ggcaactgca 28680  
ccgtctacag cgaggcgat tccctactta gtctctgttt aaccaaatgc ggagctcacg 28740  
tcttaggaag tgtaagttta accggtttaa caggaacct aaccctaatg actgatattt 28800  
ctgtcaccat tcaatttaca ttgacaaca atggtaagct actaagctct ccacttataa 28860  
acaacgcctt tagtattcga cagaatgaca gtacggctc aaacctacc tacaacgccc 28920  
tggggtttat gcctaacagt accatatatg caagaggggg aggtggtgaa ccacgaacaa 28980  
actactacgt ccaaactgat cttaggggaa atgttcaaaa accaatcatt cttactgtaa 29040  
ctacaactc agtcgccaca ggatattcct tatcttttaa gtggactgct cttgcacgtg 29100  
aaaagtgtg aacccaaca acctcgtttt gctacattac agaacaataa aacogtgtac 29160  
cccaccggtt cgttttttcc agatgaaacg ggcgagagtt gatgaagact tcaaccagtt 29220  
gtacccttat gacccccac atgctcctgt tatgcccttc attactccac cttttacctc 29280  
ctcggatggg ttgcaggaaa aaccacttgg agtgtaagt ttaaactaca gagatcccat 29340  
tactacgcaa aatgagtctc ttacaattaa actaggaaac ggcctcactc tagacaacca 29400  
gggacaacta acatcaaccg ctggcgaagt agaacctcca ctactaacg ctaacaacaa 29460  
actgcaactg gtctatagcg atccttttag agtaagcgc aacagcctaa ccttatcgca 29520  
caccgctccc cttgttattg ctgataactc ttttagcattg caagtttcag agcctatttt 29580  
tataaatgac aaggacaac tagccctgca aacagccgag ccccttgtaa ctaacgctgg 29640  
cacccttcgc ttacaagcgg ccgcccttt aggcattgca gaccaaaacc taaaactcct 29700  
gtttaccaac cttttgtact tgcagaataa ctttctcacg tttagcattg aacgaccct 29760  
tgccattacc aatactggaa agctggctct acagctctcc ccaccgctac aaacagcaga 29820  
cacaggcttg actttgcaaa ccaacgtgcc attaactgta agcaacggga ccctaggctt 29880

---

-continued

---

agccataaag cgcccactta ttattcagga caacaacttg tttttggact tcagagctcc 29940  
ctcgcgtctt ttcaacacgc acccagtact agggcttaac ttttacacc ctcttgcggt 30000  
acgcgatgag gcgctcactg ttaacacagg ccgcggcctc acagtgagtt acgatggttt 30060  
aatTTTaaat cttggttaagg atcttcgctt tgacaacaac accgTTTctg tcgctcttag 30120  
tgctgctttg cttttacaat aactgatca gcttcgcctt aacgtgggag ctgggctgag 30180  
ttacaatcca gtgagtaaga aattggacgt gaaccccaat caaacaagg gtttaacctg 30240  
ggaaatgac tacctcattg taaagctagg aaatggatta ggTTTTgatg gcgatggaaa 30300  
catagctggt tctcctcaag ttacatgcc tgacacctta tggaccactg ccgaccctc 30360  
ccccattgt tccatctaca ctgatttaga tgccaaaatg tggctctcgt tggtaaaaca 30420  
agggggtggt gttcacgggt ctggtgcttt aaaagcattg aaaggaacc tattgagtcc 30480  
tacggaaagc gccattgta ttatactaca ttttgacaat tatggagtgc gaattctcaa 30540  
ttatcccact ttgggcactc aaggcacgtt gggaaataat gcaactggg gttataggca 30600  
gggagaatct gcagacacta atgtactcaa tgcactagca tttatgccca gttcaaaaag 30660  
gtaccaaga gggcgtgaa gcgaagtca gaatcaaact gtgggctaca ctgtatata 30720  
gggtgacttt tctatgccg taccgtacca aatacagta aactatggac caactggcta 30780  
ctcctttaa tttatttga gaactgttc aagacaacca tttgacatcc catgctgttt 30840  
tttctcttac attacggaag aataaaaca ctttttctt ttttttctt tttattttac 30900  
acgcacagta aggcttcctc caccctcca tctcacagca tacaccagcc tctccccctt 30960  
catggcagta aactgttggt agtcagtcgt gtatttggga gttaagatcc aaacagtctc 31020  
tttggtagt aaacatggat ccgtgatgga cacaatccc tgggacaggt tctccaacgt 31080  
ttcggtaaaa aactgcagtc cgccctacaa aacaacagg ttcaggctct ccacggggtta 31140  
tctccccgat caaactcaga cagagtaaag gtgcgatgat gttccactaa accacgcagg 31200  
tggcgtgctc tgaacctctc ggtgcgactc ctgtgaggct ggtagaagt tagattgtcc 31260  
agcagcctca cagcatggat catcagtcta cgagtgcgctc tggcgcagca gcgcatctga 31320  
atctcactga gattccggca agaatgcac accatcacia tcagggtggt catgatccca 31380  
tagtgaaca cgctccagcc aaagctcatt cgctccaaca gcgccacgc gtgtccgtcc 31440  
aaccttactt taacataaat caggtgtctg ccgcgtacaa acatgctacc cgatacaga 31500  
acctccggg gaaaaccctt gttcaccacc tgcctgtacc agggaaaact cacatttatc 31560  
agggagccat agatagccat tttaaacaa ttagctaaca ccgccccacc agctctacac 31620  
tgaagagaac cgggagagtt acaatgacag tgaataatcc atctctcata acccctaaty 31680  
gtctgatgga aatccagatc taactgggca cagcagatc acactttcat atacattttc 31740  
atcatatggt tttcccagc cgttaaaata caatcccaat acacgggcca ctctctcagt 31800  
acaataaagc taatacaaga tggatatact ctcacctcac taacattgtg catgttcata 31860  
ttttcacatt ctaagtaccg agagttctcc tctacaacag cactgcccgc gtctcacia 31920  
gggtgtagct ggtgacgatt gtaaggagcc agtctgcagc gataccgtct gtcgcgttgc 31980  
atcgtagacc agggaccgac gcaactcctc gtactttag tagcagaacc acgtccgctg 32040  
ccagcacgtc tccaagtaac gccggtccct gcgtcgtca cgctccctcc tcaacgcaaa 32100  
gtgcaaccac tcttgaatc cacacagatc cctctcgcc tccggggoga tgcacacctc 32160

---

-continued

---

aaacctacag atgtctcggg acagttccaa acacgtagtg agggcgagtt ccaaccaaga 32220  
cagacagcct gatctatccc gacacactgg aggtggagga agacacggaa gaggcattgt 32280  
attccaagcg attcaccacac gggtcgaaat gaagatcccg aagatgacaa cggtcgcctc 32340  
cggagccctg atggaattta acagccagat caaacattat gcgattttcc aggctatcaa 32400  
tcgcgccctc caaaagagcc tggaccgcga cttccacaaa caccagcaaa gcaaaagcgt 32460  
tattatcaaa ctcttcgatc atcaagctgc aggactgtac aatgcccaag taattttcat 32520  
ttctccactc gcgaatgatg tcgcgcaaaa tagtctgaag gttcatgccg tgcatattaa 32580  
aaagctccga aagggcgccc tctatagcca tgcgtagaca caccatcatg actgcaagat 32640  
atcgggctcc tgagacacct gcagcagatt taacagacc aggtcaggtt gctctccgcg 32700  
atcgcgaatc tccatccgca aagtcatttg caaataatta aatagatctg cgcgcactaa 32760  
atctgttaac tccgcgctag gaactaaatc aggtgtggct acgcagcaca aaagttccag 32820  
ggatggcgcc aaactcacta gaaccgctcc cgagtagcaa aactgatgaa tgggagtaac 32880  
acagtgtaaa atgttcagcc aaaaatcact aagctgctcc tttaaaagt ccagtacttc 32940  
tatattcagt tcgtgcaagt actgaagcaa ctgtgcgggg atatgcacag caaaaaaat 33000  
agggcgctc agatacatgt tgacctaaaa taaaagaat cattaaacta aagaagcctg 33060  
gcgaacgggt ggatatatga cacgctccag cagcaggcaa gcaaccggct gtccccggga 33120  
accgcggtaa aattcatccg aatgattaa aagaacaaca gagacttccc accatgtact 33180  
cggttggatc tcctgagcac agagcaatc cccctcaca tcatatccg ctacagaaaa 33240  
aaaaagctcc agataccag cgggaatc caacgacagc tgcaagaca gcaaaacaat 33300  
ccctctggga gcaatcacia aatctcccg tgaaaaagc acatacatat tagaataacc 33360  
ctgttgctgg ggcaaaaag cccgctgctc cagcaaatgc acataaatat gttcatcagc 33420  
cattgccccg tcttaccgcg taaacagcca cgaaaaatc gagctaaaat ccaccaaca 33480  
gcctatagct atatatcac tccaccaat gagcctaata ccgcaccacc cagaccaaa 33540  
gttcaccac acccaciaaa cccgcgaaaa tccagcgcg tcagcacttc cgcaatttca 33600  
gtctcaaac gtcacttccg cgcgcctttt cactttcca cacacgcct tcgcccgcc 33660  
gccctcgcgc cccccgcgt cccccacgt caccgcacgt cccccggcc ccgctcgt 33720  
cctccccgct cattatcata ttggcacggt tccagaataa ggtatattat tgatgcagca 33780  
aaacaatccc tctgggagca atcacaaaat cctccggtga aaaaagcaca tacatattag 33840  
aataaccctg ttgtggggc aaaaaggccc gtcgtcccag caaatgcaca taaatatgtt 33900  
catcagccat tgcccgtct taccgctaa acagccacga aaaaatcgag ctaaaatcca 33960  
cccaacagcc tatagctata tataactcc acccaatgac gctaataccg caccaccac 34020  
gaccaaagt caccacacc cacaaaacc gogaaaatcc agcgcgctca gcaottccgc 34080  
aatttcagtc tcacaacgct acttcgcgc gccttttcac tttccacac acgcccctcg 34140  
cccgcgcgc ctgcgcacc cccgcgtcac cccacgtcac cgcacgtcac cccggccccg 34200  
cctcgtcct cccgcgtcat tatcatattg gcacgtttcc agaataaggt atattattga 34260  
tgca 34264

---

-continued

---

<212> TYPE: DNA

<213> ORGANISM: simian adenovirus SV-25

<400> SEQUENCE: 5

catcatcaat aatatacctt attctgaaaa cgtgccaata tgataatgag cggggaggag 60  
cgaggcgggg ccgggggtgac gtgcgggtgac gcgggggtggc gcgagggcgg ggcgaagggc 120  
gcgggtgtgt gtgtgggagg cgcttagttt ttacgtatgc ggaaggagggt tttataccgg 180  
aagatgggta atttgggcgt atacttgtaa gttttgtgta atttggcgcg aaaactgggt 240  
aatgaggaag ttgaggttaa tatgtacttt ttatgactgg gcggaatttc tgctgatcag 300  
cagtgaactt tgggcgctga cggggagggt tcgctactg acagtaccac gagaaggctc 360  
aaaggccca tttattgtac tcttcagcgt tttcgctggg tatttaaacg ctgtcagatc 420  
atcaagaggc cactcttgag tgctggcgag aagagttttc tcctccgtgc tgccacgatg 480  
aggctggtcc ccgagatgta cgggtgtttt agcgacgaga cggtgcgtaa ctcagatgac 540  
ctgctgaatt cagacgcgct ggaaatttcc aattcgcctg tgctttcggc gccgtcactt 600  
cacgacctgt ttgtgttttg gctcaacgct tagcaacgtg ttatataggg tcaagaagga 660  
gcaggagacg cagtttgcta ggctgttggc cgatactcct ggagtttttg tggctctgga 720  
tctaggccat cactctcttt tccaagagaa aattatcaa aacttaactt ttacgtctcc 780  
tggtcgcacg gttgcttccg ctgcctttat tacctatatt ttggatcaat ggagcaacag 840  
cgacagccac ctgtcgtggg agtacatgct ggattacatg tcgatggcgc tgtggagggc 900  
catgctgagg aggagggttt gcatttactt gcgggcgcag cctccgcggc tggaccgagt 960  
ggaggaggag gacgagccgg gggagaccga gaacctgagg gccgggctgg acctccaac 1020  
ggaggactag gtgctgagga tgatcccga gaggggacta gtggggctag gaagaagcaa 1080  
aagactgagt ctgaacctcg aaactttttg aatgagttga ctgtgagttt gatgaatcgt 1140  
cagcgtccgg agacaatttt ctggtctgaa ttggaggagg aattcaggag gggggaactg 1200  
aacctgctat acaagatgag gtttgaacag ttaaaaactc actggttggg gccgtgggag 1260  
gattttgaaa ccgccttgga cacttttgct aaagtggctc tgcggccgga taaggtttac 1320  
actatccgcc gactgttaa cataaagaag agtgtttatg ttataggcca tggagctctg 1380  
gtgcagggtc aaaccgtcga ccgggtggcc tttagttgag gtatgcaaaa tctgggcccc 1440  
ggggtgatag gcttaaatgg tgtaacattt cacaatgtaa ggtttactgg tgaagtttt 1500  
aacggctctg tgtttgcaaa taacacacag ctgacgctcc acggcgttta cttttttaac 1560  
tttaataaca catgtgtgga gtcgtggggc aggggtgctt tgaggggctg ctgttttcac 1620  
ggctgctgga agcgggtggt gggaaagactt aaaagtgtaa catctgtaaa aaaatgcgtg 1680  
tttgagcggg gtgtgttggc ttaactgtg gagggctgtg gacgcattag gaataatgcg 1740  
gcgtctgaga atggatgttt tcttttgcta aaaggcacgg ctagtattaa gcataacatg 1800  
atatcggcca gcggtctgta cccttcacag ctgttaactt gcgcggatgg aaactgtcag 1860  
accttgcgca ccgtgcacat agcgtccac cagcgcggc cctggccaac attcagcac 1920  
aatatgctta tgcgttggc cgtccacttg ggccttaggc gaggcgtgtt tgtgccttac 1980  
cagttaact ttagccatac caagatttta ctagaacctg ataccttctc tcgagtgtgt 2040  
ttcaatgggg tgtttgacat gtcaatggaa ctgtttaaag tgataagata tgatgaatcc 2100  
aagtctcgtt gtcgcccacg tgaatgcgga gctaatcatc tgaggttgta tcctgtaacc 2160

-continued

---

|            |            |             |            |             |            |      |
|------------|------------|-------------|------------|-------------|------------|------|
| ctaaacgtta | ccgaggagct | gaggacggat  | caccacatgt | tgtcctgcct  | gcgcaccgac | 2220 |
| tatgaatcca | gcgacgagga | gtgaggtgag  | gggcggagcc | acaaagggta  | taaaggggcg | 2280 |
| tgaggggtgg | gtgtgatgat | tcaaaatgag  | cgggacgacg | gacggcaacg  | cgtttgaggg | 2340 |
| tggagtgttc | agcccttata | tgacatctcg  | tcttccttcc | tgggcaggag  | tgcgtcagaa | 2400 |
| tgtagtgggc | tccaccgtgg | acggaacgacc | ggtcgcccct | gcaaatccg   | ccaccctcac | 2460 |
| ctatgccacc | gtgggatcat | cgttggacac  | tgccgcggca | gctgccgett  | ctgetgccgc | 2520 |
| ttctactgct | cgcggcatgg | cggctgattt  | tggactgtat | aaccaactgg  | ccactgcagc | 2580 |
| tgtggcgtct | cggctctcgg | ttcaagaaga  | tgccctgaat | gtgatcctga  | ctcgcctgga | 2640 |
| gatcatgtca | cgtcgttgg  | acgaaactggc | tgccgagata | tcccaageta  | accccatac  | 2700 |
| cacttcagaa | tccataaata | aagacaaaca  | aatatgttga | aaagtaaaat  | ggctttat   | 2760 |
| gtttttttt  | gctcggtagg | ctcgggtcca  | cctgtctcgg | tcgttaagaa  | ctttgtgtat | 2820 |
| gttttccaaa | acacggtaca | gatgggcttg  | gatgttcaag | tacatgggca  | tgaggccatc | 2880 |
| tttgggtga  | agataggacc | attgaagagc  | gtcatgctcc | ggggtggtgt  | tgtaaattac | 2940 |
| ccagtcgtag | cagggtttct | ggcgtggaa   | ctggaagatg | tcctttagga  | gtaggctgat | 3000 |
| ggccaagggc | aggcccttag | tgtaggtgtt  | tacaaagcgg | ttaagctggg  | agggatgcat | 3060 |
| gcggggggag | atgatatgca | tcttggcttg  | gatcttgagg | ttagctatgt  | taccaccag  | 3120 |
| gtctctcggg | gggttcatgt | tatgaaggac  | caccagcacg | gtgtagccgg  | tgcatttggg | 3180 |
| gaacttgtca | tgcatgttgg | aggggaagc   | gtggaagaat | ttagagacc   | ccttgtggcc | 3240 |
| ccctaggttt | tccatgcact | catccataat  | gatggcaatg | ggaccctgg   | cggccgcttt | 3300 |
| ggcaaacacg | ttttgggggt | tggaaacatc  | atagttttgc | tctagagtga  | gctcatcata | 3360 |
| ggccatctta | acaaagcggg | gtaggagggt  | gcccgactgg | gggatgatag  | ttccatctgg | 3420 |
| gcctggggcg | tagttaccct | cacagatctg  | catctcccag | gccttaattt  | cagagggggg | 3480 |
| tatcatgtcc | acctgggggg | caataaagaa  | cacggtttct | ggcgggggat  | tgatgagctg | 3540 |
| ggtgaaagc  | aagttacgca | gcagttgaga  | tttgccacag | ccggtggggc  | cgtagatgac | 3600 |
| cccgatgacg | ggttgacgct | gtagttgag   | agaggaacag | ctgccgtcgg  | ggcgcaggag | 3660 |
| gggggctacc | tattcatca  | tgttctaac   | atgtttat   | tcaactcacta | agttttgcaa | 3720 |
| gagcctctcc | ccaccaggg  | ataagagttc  | ttccaggctg | tgaagtgtt   | tcagcggttt | 3780 |
| taggccgtcg | gcatgaggca | tcttttcgag  | cgactgacga | agcaagtaca  | gtcggtecca | 3840 |
| gagctcgggt | acgtgctcta | tggaatctcg  | atccagcaga | cttcttgggt  | gcgggggtt  | 3900 |
| ggtcgacttt | cgtgtaggg  | caccagccgg  | tggcgtcca  | ggccgcgag   | ggttctgtcc | 3960 |
| ttccagggtc | tcagcgtccg | ggtgagggtg  | gtctcgggtg | cgggaaagg   | atgagccccg | 4020 |
| ggctggggcg | ttgcgaggt  | gcgttcag    | ctcatcctgc | tgggtctgaa  | gcggacgtcg | 4080 |
| tctccctgtg | agtcggccag | atagcaacga  | agcatgaggt | cgtagctgag  | ggactcggcc | 4140 |
| gcgtgtccct | tggcgcgag  | ctttcccttg  | gaaacgtgct | gacatttgg   | gcagtgcaga | 4200 |
| cattggaggg | cgtagagttt | ggggccag    | aagaccgact | cgggcgagta  | ggcgtcggct | 4260 |
| ccgactgag  | cgcagacggt | ctcgcactcc  | actagccacg | tgagctcggg  | tttagcggga | 4320 |
| tcaaaaacca | agttgcctcc | atttttttt   | atgcgtttct | tacctgctg   | ttccatgagt | 4380 |
| ttgtggcccc | cttccgtgac | aaaaaggctg  | tcggtgtctc | cgtagacaga  | cttgaggggg | 4440 |

---

-continued

---

cgatcttcca aagggtttcc gaggtcttcc gcgtacagga actgggacca ctccgagacg 4500  
aaggctctcg tccaggetaa cacgaaggag gcaatctcgg aggggtatct gtcgttttca 4560  
atgagggggt ccaccttttc caggggtgtc agacacaggt cgtcctcctc cgcgtccacg 4620  
aagggtgattg gcttghtaagt gtaggtcacg tgatctgcac cccccaagg ggtataaaag 4680  
ggggcgtgcc caccctctcc gtcactttct tccgcatcgc tgtggaccag agccagctgt 4740  
tcgggtgagt aggccctctc aaaagccggc atgatctcgg cgtcaagtt gtcagtttct 4800  
acaaacgagg tggatttgat attcacgtgc cccgcggcga tgcttttgat ggtggagggg 4860  
tccatctgat cagaaaaac gatctttttg ttgtcaagtt tggggcgaa agaccgtag 4920  
agggcgttgg aaagcaactt ggcgatggag cgcagggctc gatttttctc ccgatcggcc 4980  
ctctccttgg cggcgtgatt gagttgcacg tactccccgg cgcgcaccg cactcgggg 5040  
aacacggcgg tgcgctcgtc gggcaggatg cgcacgcgcc agccgcgatt gtgcagggtg 5100  
atgaggtcca cgctggtagc cacctccccg cggaggggct cgttgggtcca acacaatcgc 5160  
cccccttttc tggagcagaa cggaggcagg ggatctagca agttggcggg cggggggctg 5220  
gcgtcgtagg tgaagatacc gggtagcagg atcttattaa aataatcgtat ttcggtgtcc 5280  
gtgtcttgca acgcgtcttc ccaacttctc accgccaggg ccctttcgtg gggattcagg 5340  
ggcgtcccc agggcatggg gtgggtcagg gccgaggcgt acatgccgca gatgtcatac 5400  
acgtacaggg gttccctcaa cccccgatg taagtggggg aacagcggcc cccgcggatg 5460  
ctggctcgca cgtagtcgta catctcgcgc gagggagcca tgaggccgct tcccaagtgg 5520  
gtcttgtggg gtttttcggc ccggtagagg atctgtctga agatggcgtg ggagttggaa 5580  
gagatggtgg ggcgttgaa gacgttaaag ttggccccgg gtagtccac ggagtcttg 5640  
atgaactggg cgtaggatcc ccggagttt tccaccaggg cggcggtcac cagcacgtcg 5700  
agagcgcagt agtccaaact ctcgcggacc aggtttagg ccgtctctt tttttctcc 5760  
cacagttcgc ggttgaggag gtattctcgc cggcttttcc agtactcttc ggcgggaaat 5820  
cctttttcgt ccgctcggtg agaacctaac atgtaaaatt cgttcaccgc tttgatgga 5880  
caacagcctt tttctaccg cagggcgtac gcttgagcgg cctttctgag agaggtgtgg 5940  
gtgagggcga aggtgtccc caccatcact ttcaggtact gatgtttgaa gtcogtgcg 6000  
tcgacggcgc cctgttccca cagcgtgaa tccgtgcgct tttctgcct gggattgggg 6060  
agggcgaagg tgacatcgtt aaagagtatt ttcccggcgc gggcatgaa gttgcgagag 6120  
atctgaagg gcccgggac gtcgagcgg ttgttgatga cctgcgcgc caggacgatc 6180  
tcgtcgaagc cgttgatgtt gtgaccacg atgtaaagtt cgtgaagcg cggctgtccc 6240  
ttgagggcgc gcgctttttt caactcctcg taggtgagac agtccggcga ggagagacce 6300  
agctcagccc gggcccagtc ggagagttga ggattagccg caaggaagga gctccataga 6360  
tccaaggcca ggagagtttg caagcggtcg cggaaactcgc ggaacttttt cccacggcc 6420  
attttctccg gtgtcactac gtaaaaggtg ttggggcggg tgttccacac gtcocatcgg 6480  
agctctaggg ccagctcgca ggcttggcga acgagggctc cctcgcaga gacgtgcatg 6540  
accagcataa aggttaccaa ctgtttcccg aacgagccca tccatgtgta ggtttctacg 6600  
tcgtagggtg caaagagcgg ctgggtgcgc gcgtgggagc cgtcggaaa gaagctgatc 6660  
tcctgccacc agctggagga atgggtgtta atgtggtgga agtagaagtc ccgcccggc 6720

---

-continued

---

acagagcatt cgtgctgatg tttgtaaaag cgaccgcagt agtcgcagcg ctgcacgctc 6780  
tgtatctcct gaacgagatg cgcttttcgc ccgcgcacca gaaaccggag ggggaagtgt 6840  
agacgggggg ctggtggggc gacatcccct tcgccttggc ggtgggagtc tgcgtctgcg 6900  
tcctccttct ctgggtggac gacggtgggg acgacgacgc cccgggtgcc gcaagtccag 6960  
atctccgcca cggaggggtg cagggcgtgc aggaggggac gcagctgccc gctgtccagg 7020  
gagtcgaggg aagtcgcgct gaggtcggcg ggaagcgttt gcaagttcac tttcagaaga 7080  
ccggtaagag cgtgagccag gtgcagatgg tacttgattt ccaggggggt gttggatgaa 7140  
gcgtccacgg cgtagaggag tccgtgtccg cgcggggcca ccaccgtgcc ccgaggaggt 7200  
tttatctcac tcgtcagagg cgagcggcgg ggggtagagg cggctctgcg ccggggggca 7260  
gcggaggcag aggcacgttt tcgtgaggat tcggcagcgg ttgatgacga gcccgagac 7320  
tgctggcgtg ggcgacgac cgcgcggtga ggtcctggat gtgccgtctc tgcgtgaaga 7380  
ccaccggccc ccgggtcctg aacctaaaga gagttccaca gaatcaatgt ctgcatcgtt 7440  
aacggcgccc tcctcagga tctcctgcac gtcgcccag ttgtcctgat aggcgatctc 7500  
ggccatgaac tgttccactt ctctctcgc gaggtcaccg tggcccgtc gctccacggt 7560  
ggcggccagg tcgttgaga tgcggcgcag gagttgagag aaggcgttga ggcggttctc 7620  
gttccacacg cggtgtaca ccacgtttcc gaaggagtgc cgcctcgcga tgaccacctg 7680  
ggccacgttg agttccactt ggcggcgaa gacggcgtag tttctgaggc gctggaagag 7740  
gtagttgagc gtggtggcga tgtgctgcga gacgaagaag tacataatcc agcggccgag 7800  
ggtcatctcg ttgatgtctc cgtatgcttc gagacgtcc atggcctcgt agaagtgcac 7860  
ggcgaagtgt aaaaattggg agttgcgggc ggccaccgtg agttcttctt gcaggaggcg 7920  
gatgagatcg gcgaccgtgt cgcgcacctc ctgttcgaaa gcgcccagag gcgcctctgc 7980  
ttcttctcgc ggctcctct ctccagggg ctcgggttcc tccggcagct ctgcgacggg 8040  
gacggggcgg cgacgtcgtc gctcgtaccg cagggcgtcc acgaagcgtc cgtatcttc 8100  
gccgcgccgg cgacgatgg tctcgggtgac ggcgcgtccg ttttcgcgag gtcgcagttc 8160  
gaagacgccg ccgcgcagag cgcgccctg cagggagggt aagtggttag ggcgctcggg 8220  
cagggacacg gcgctgacga tgcattttat caattgctgc gtaggcactc cgtgcagggg 8280  
tctgagaacg tcgaggtcga cgggatccga gaacttctct aggaaagcgt ctatccaatc 8340  
gcaatcgcga ggtaagctga gaacggtggg tcgctggggg gcgttcgagg gcagttggga 8400  
ggtgatgctg ctgatgatgt aattaaagta ggcgtcttc aggcggcggg tgggtggcag 8460  
gaggaccacg tctttgggcc cggcctgttg aatgcgcagg cgctcggcca tgccccaggc 8520  
ctcgtctga cagcgacgca ggtctttgta gaagtcttgc atcagtctct ccaccggaac 8580  
ctctgcttct ccctgtctg ccatgogagt cgagccgaac ccccgaggg gctgcagcaa 8640  
cgctaggtcg gccacgacct tttcggccag cacggcctgt tgaatctgag tgagggtggc 8700  
ctggaagtgc tccaggtcca cgaagcgggt ataggcccc gtgttgatgg ttaggtgca 8760  
gttgccatg acggaccagt tgacgacttg catgccgggt tgggtgatct ccgtgtactt 8820  
gaggcgcgag taggcctcgg actcgaacac gtagtcgttg catgtgcgca ccagatactg 8880  
gtagccgacc aggaagtgag gaggcggctc tcggtacagg ggcagccaa cggtggcggg 8940  
ggcggcgggg gacaggtcgt ccagcatgag gcggtggtag tggtagatgt agcgggagag 9000



---

-continued

---

ccaggtgatg ccggccgagg tggttgcgpc cctggtgaat tcgcgacgc ggttccagat 9060  
gttgccgagg ggaccaaagc gctccatggt gggcacgctc tgccccgtga ggcgggpcga 9120  
atcttgtagc ctctagatgg aaaaagaca gggcggcatc cgactccttt ccgtagcttg 9180  
gggggtaaaag tcgcaagggt gcggcggcgg ggaaccccg ttcgagaccg gccggatccg 9240  
ccgctcccga tcgacctggc cccgcatcca cgacgtccgc gccgagacc agccgcgacg 9300  
ctccgcccc atacggaggg gagtcttttg gtgttttttc gtagatgcat ccggtgctgc 9360  
ggcagatgag accccagagc cccactacca ccgccgtggc ggcagtaaac ctgagcggag 9420  
gctgtgacag ggaggaggaa gactgtgctt tagacctgga agaggagag gggctggccc 9480  
ggctgggagc gccatcccc gagagacacc ctagggttca gctcgtgagg gacgccaggc 9540  
aggcttttgt gccgaagcag aacctgttta gggaccgag cggtcaggag gccggaggaga 9600  
tcgctgattg caggtttcgg gcgggagag agctcagggc gggcttcgat cgggagcggc 9660  
tcctgagggc ggaggatttc gagccgacg agcgttctgg ggtgagccc gcccgctc 9720  
acgtatcggc ggccaacctg gtgagcgcgt acgagcagc ggtgaacgag gagcgcaact 9780  
tccaaaagag cttaacaat cacgtgagga cctgatcgc gagggaggag gtgacctcg 9840  
ggctgatgca tctgtgggac ttcgtggagg cctacgtgca gaacccggct agcaaacc 9900  
tgacggcccc gctgttccgt atcgtgacg acagccgca caacgagacg ttcgagcag 9960  
ccatgttgaa catcgcggag cccgagggtc gctggctctt ggatctgatt aacatcctgc 10020  
agagcatcgt ggtgcaggag agggcctga gtttagcgga caagggtggc gccattaact 10080  
attcagatgca gagcctgggg aagttctacg ctcgcaagat ctacaagagc ccttacgtgc 10140  
ccatagacaa ggaggtaag atagacagct tttacatgag catggcctg aagggtctga 10200  
cgctgagcga cgacctcggc gtgtaccgta acgacaagat ccacaaggcg gtgagccca 10260  
gccgcccggc ggagctgagc gacagggagc tgatgcacag cctgcagagg gcgctggcgg 10320  
gcgcccggga cgaggagcgc gaggcttact tcgacatggg agccgatctg cagtggcgtc 10380  
ccagcgcgcg cgcttgagg gcggcgggtt atcccgacga ggaggatcgg gacgatttg 10440  
aggaggcagg cgagtacgag gacgaagcct gaccgggag gtgttgtttt agatgcagcg 10500  
gccggcggac gggaccaccg cggatcccgc acttttgca tccatgcaga gtcaccttc 10560  
gggctgacc gcctccgatg actggcggc ggccatggac cgcatcatgg cgctgaccac 10620  
ccgcaacccc gaggtttta gccagcaacc ccaggccaac cgtttttcgg ccatcttga 10680  
agcggtggtg ccgtcgcga ccaaccgac gcacgagaaa gtcctgacta tcgtgaacgc 10740  
cctggtagac agcaaggcca tccgccgtga cgaggcggc ttgatttaca acgctctttt 10800  
ggaaacgctg gcgctgata acagcactaa cgtgcagacc aatctggacc gcctaccac 10860  
cgacgtgag gaggcctgg gcgagaagga goggtttctg agggacagta atctgggctc 10920  
tctggtggca ctgaacgcct tcctgagctc acagccggcc aacgtgccc gcgggagga 10980  
ggattacgtg agcttcatca gcgctctgag actgctggtg tccgaggtgc cccagagcga 11040  
ggtgtaccag tctggccgg attacttttt ccagacgtcc cgacagggt tgcaaaccgt 11100  
gaacctgact caggccttta aaaactgca aggcattgg ggggtcaagg ccccggtgg 11160  
cgatcgcgc actatctcca gtctgctgac cccaacact cgctgctgc tgetcttgat 11220  
cgcaccgtt accaacagta gcaactatcag ccgtgactcg tacctgggtc atctcatcac 11280

---

-continued

---

tctgtaccgc gaggccatcg gccaggctca gatcgacgag catacgtatc aggagattac 11340  
taacgtgagc cgtgccctgg gtcaggaaga taccggcagc ctggaagcca cgttgaactt 11400  
tttgctaacc aaccggaggc aaaaaatacc ctcccagttc acgttaagcg ccgaggagga 11460  
gaggattctg cgatacgtgc agcagtcctg gagcctgtac ttgatgcgcg agggcgccac 11520  
cgcttccacg gctttagaca tgacggctcg gaacatggaa ccgtcctttt actccgcccc 11580  
ccggccggtc attaacctgc tgatggacta cttccatcgc gcggccgcca tgaacgggga 11640  
gtacttcacc aatgccatcc tgaatccgca ttggatgccc ccgtccggct tctacaccgg 11700  
ggagtttgac ctgcccgaag ccgacgacgg ctttctgtgg gacgacgtgt ccgatagcat 11760  
ttcacgccc gctaactcgc gattccagaa gaaggagggc ggagacgagc tccccctctc 11820  
cagcgtggaa gcggcctcaa ggggagagag tccctttcca agtctgtctt ccgccagtag 11880  
cggtcgggta acgctccac ggttgccggg ggagagcgac tacctgaacg accccttgct 11940  
gcgaccggct agaaagaaaa attttcccaa taacggggtg gaaagcttg tgataaaat 12000  
gaatcgttg aagacgtacg ccagggagca gcgggagtgg gaggacagtc agccgcgcc 12060  
gctggtaccg ccgcatggc gtcgcccag agaaagaccg gacgactccg cagacgatag 12120  
tagcgtggtg gacctgggag ggagcggagc caacccttt gctcactgc aaccaaggg 12180  
gcgctcagct cgcctgtatt aataaaaaag acgcggaac ttaccagagc catggccaca 12240  
gcgtgtgtgc tttcttctc tcttcttcc tggcgcggc agaatgagaa gagcggtgag 12300  
agtcacgccc gcggcgtatg agggcccgc cccttcttac gaaagcgtga tgggatcagc 12360  
gaacgtgccc gccacgtgg agggcctta cgttctctcc agatacctgg gacctacgga 12420  
gggcagaaac agcatccgtt actccgagct ggcgccctg tacgatacca ccaaggtgta 12480  
cctggtggac aacaagtcgg cggacatcgc ctccctgaat taccaaaacg atcacagtaa 12540  
ctttctgact accgtggtgc agaacaatga cttcaccctc acggaggcgg gcacgcagac 12600  
cattaacttt gacgagcgtt cccgctgggg cggtcagctg aaaaccatcc tgcacaccaa 12660  
catgcccac atcaacgagt tcatgtccac caacaagttc agggctaagc tgatggtaga 12720  
aaaaagtaat gcggaactc ggcagcccg atacgagtgg ttcgagtta ccattccaga 12780  
gggcaactat tccgaaacta tgactatcga tctcatgaat aacgcgatc tggacaatta 12840  
cctgcaagtg ggagacaga acggggtgct ggaaagcgt atcggcgtga aattcgatac 12900  
cagaaacttc cgactggggt gggatcccgt gaccaagctg gtgatgccag gcgtgtacac 12960  
caacgaggct tttcaccctg acatcgtgct gctgccggg tgcggtgtgg acttactca 13020  
gagccgtttg agtaacctgt taggaattag aaagcgcgc cccttccaag agggctttca 13080  
aatcatgtat gaggacctg agggaggtaa tatacccgcc ttactggagc tgcgaagta 13140  
cgaagctagc atacaacgcg ccaaagcggg gggtagagag attcggggag acaccttgc 13200  
ggtagctccc caggacctg aatatgtgc tttactaaa gacagcaaag acagaagcta 13260  
caatattata aacaacacga cggacacct gtatcggagc tggtttctg cttacaacta 13320  
cggagacccc gagaaaggag tgagatcatg gaccatactc accaccaagg acgtgacctg 13380  
tggtcgcag caagtgtact ggtccctgcc ggatgatg caagaccgg tcacctccg 13440  
ccccccacc caagtacga acttcccgtt ggtgggacc gagctgctgc ccgtccatgc 13500  
caagagcttc tacaacgagc aggccgtcta ctgcgaactt attcgcagc ccaaccgct 13560

---

-continued

---

taccacgctg ttcaatcgct ttcccagaaa ccagattctg gtgcgccttc ccgctcctac 13620  
cattaccacc gtcagtgaaa acgttcccgc cctcacagat cacggaaccc tgccgctgcg 13680  
cagcagtatc agtggagttc agcgcgtgac catcaccgac gccagacgtc gaacctgccc 13740  
ctacgtttac aaagcgcttg gcgtgggtgc tcctaaagtt ctttctagtc gcaccttcta 13800  
aaaaatgctc catcctcatc tctcccagata acaacaccgg ctggggactg ggctccggca 13860  
agatgtacgg cggagcctaa aggcgctcca gtcagcacc agttcgagtt cggggccact 13920  
tccgctctcc ttggggagct tacaagcgag gactctcggg tcgaacggct gtagacgata 13980  
ccatagatgc cgtgattgcc gacgcccgcc ggtacaaccc cggaccggtc gctagcggcg 14040  
cctccaccgt ggattccgtg atcgacagcg tggtagccgg cgctcgggcc tatgctcgcc 14100  
gcaagaggcg gctgcatcgg agacgctgcc ccaccgccgc catgctggca gccaggggcg 14160  
tgctgagggc gggccggagg gcaggcagaa gggctatgcg ccgctgctgc gccaacggcg 14220  
ccgcccggag gggcccga caggctgccc gccaggctgc cgctgccatc gctagcatgg 14280  
ccagaccag gagaggaac gtgtactggg tgcgtgattc tgtgacggga gtcgagtgct 14340  
cggctgagcg ccgacctccc cgaagttaga agatccaagc tcggaagacg gcggtactga 14400  
gtctccctgt tgttatcagc ccaacatgag caagcgcaag tttaaagaag aactgctgca 14460  
gacgctgggt cctgagatct atggccctcc ggacgtgaag ccagacatta agccccgca 14520  
tatcaagcgt gttaaaaagc gggaaaagaa agaggaactc gcggtggtag acgatggcgg 14580  
agtgaattt attagaggtt tcgccccgcg acgcagggtt caatggaaaag ggcggcgggt 14640  
acaacgcgtt ttgagggcgg gcaccgcggt agtttttacc ccgggagagc ggtcggccgt 14700  
taggggttcc aaaagcgagt acgacgaggt gtacggcgac gaggacatat tggaacaggc 14760  
ggctcaacag atcggagaat ttgcctacgg aaagcgttcg cgctcgcaag acctggccat 14820  
cgcttagac agcggcaacc ccacgcccag cctcaaaccc gtgacgctgc agcagggtgt 14880  
tcccgtagc gccagcagc acagcaagag ggggattaag agagaaatgg aagatctgca 14940  
tcccaccatc caactcatgg tccctaaacg gcagaggctg gaagaggctc tggagaagat 15000  
gaaagtggac cccagcatag agcggatgt aaaagtcaga cctattaagg aagtggcccc 15060  
cggctctggg gtgcaaacgg tggacattca aatccccgtc accaccgctt caaccgctg 15120  
ggaagctatg gaaacgcaaa cggagacccc tgccgcgac ggtaccaggg aagtggcgtt 15180  
gcaaacggag ccttggtacg aatacgcagc ccctcggcgt cagaggcgtt ccgctcgcta 15240  
cggccccgcc aacgccatca tgccagaata tgcgctgcat ccgtctattc tgcccactcc 15300  
cggataccgg ggtgtgacgt atcgcctgct tggaaaccgc cggcgaaccc gtcgcccgg 15360  
ccgctcccg cgcgctctgg cccccgtgct ggtgcggcgt gtgaccgccc ggggaaagac 15420  
agtgtcatt cccaaccgcg gttaccacc tagcatcctt taataactct gccgttttgc 15480  
agatggctct gacttgcccg gtgcgccttc cgttccgca ctatcgagga agatctcgtc 15540  
gtaggagagg catgacgggc agtggctgcc ggcgggcttt gcgcaggcgc atgaaaggcg 15600  
gaatthtacc cggcctgata ccataaattg ccgcccctat cgggtccata cccggcgttg 15660  
cttcagtggc gttgcaagca gctcgttaata aataaaciaa ggcttttgca cttatgacct 15720  
ggtcctgact atthttatgca gaaagagcat ggaagacatc aatthttacgt cgtggtctcc 15780  
gcggcacggc tcgcccgcgc tcatggggac ctggaacgac atcggcacca gtcagctcaa 15840

---

-continued

---

cgggggcgct ttcaattggg ggagcctttg gagcggcatt aaaaactttg gctccacgat 15900  
taaatcctac ggcagcaaa cctggaacag tagtgctggt cagatgctcc gagataaact 15960  
gaaggacacc aacttccaag aaaaagtggg caatgggggtg gtgaccggca tccacggcgc 16020  
ggtagatctc gccaaccaag cgggtgcagaa agagattgac aggcgttttg aaagctcgcg 16080  
ggtgcccgcc cagagagggg atgaggtgga ggtcggaggaa gtagaagtag aggaaaagct 16140  
gcccccgctg gagaaagtcc ccggtgcgcc tccgagaccg cagaagcggc ccaggccaga 16200  
actagaagag actctggtga cggagagcaa ggagcctccc tcgtacgagc aagccttga 16260  
agagggcgcc tctccaccct cctaccgat gactaagccg atcgcacca tggtcgcacc 16320  
ggtgtacggc aaggattaca agcccgtcac gctagagctg cccccaccgc cccccacgcg 16380  
cccagccgtc cccccctgc cgaactccgc ggccggccgc gccggaccgc tgtccgcacc 16440  
atccgctgtg cctctgccag ccgccctcc agtggccgtg gccactgcca gaaacccag 16500  
aggccagaga ggagccaact ggcaaacac gctgaacagc atcgtgggcc tgggagtga 16560  
aagcctgaaa cgcgcctgtt gctattatta aaaaagtga gctaaaaagt ctcccgttgt 16620  
atacgcctcc tatgttaccg ccagagacga gtgactgtcg ccgcgagcgc cgttttcaag 16680  
atggccacc catcgatgat gccgcagtgg tcttacctgc acatcgccgg ccaggacgcc 16740  
tcggagtacc tgagtcccgc cctcgtcag tttgcccgc ccaccgacac ctacttcagc 16800  
ttgggaaaca agtttagaaa ccccaccgtg gccccacc acgatgtgac cacggaccgc 16860  
tcgcagaggc tgaccctgc ctttgtgccc gtagaccggg aggacaccgc gtactcttac 16920  
aaagtgcgct acacgttggc cgtaggggac aaccgagtgc tggacatggc cagcacctac 16980  
ttgacatcc ggggggtgct ggatcggggg cccagcttca agccctattc cggcaccgct 17040  
tacaactccc tggccccaa gggagctccc aaccctcgg aatggacgga cacttccgac 17100  
aacaactta aagcatatgc tcaggctccc taccagagtc aaggacttac aaaggatggt 17160  
attcaggttg ggctagtgtg gacagagtca ggacaaacac cccaatatgc aaacaaagt 17220  
taccaccccg agccacaat tggggaaaac caatggaatt tagaacaaga agataaagcg 17280  
gcggaagag tcctaagaa agatacccct atgtttccct gctatgggtc atatgccagg 17340  
ccccaaaacg aacaaggagg gcaggcaaaa aaccaagaag tagatttaca gttttttgcc 17400  
actccgggcy acaccagaa cacggctaaa gtgtacttt atgctgaaa tgcacactg 17460  
gaaactccag atactcactt agtgtttaa cccgatgac acagcaccag ttcaaaactt 17520  
cttcttgggc agcaggctgc acctaacaga cccaactaca taggttttag agataatttt 17580  
attggtttaa tgtactacaa tagcactgga aacatgggcy tgctggccgg acaggcttct 17640  
caattgaat ccgtagtcga cttgcaggac agaaacaccg agttgtccta ccagctgatg 17700  
ctggacgcac tgggggatcg cagccgatat ttttcaatgt ggaatcaggc agtagacagc 17760  
tatgaccag acgttagaat tatagaaaac cacggagtgg aagacgaact gccaaactat 17820  
tgttttcctc tgggaggaat ggtggtgact gacaattaca actctgtgac gcctcaaaat 17880  
ggaggcagt gaaatacatg gcaggcagac aatactacat ttagtcaaag aggagcgcag 17940  
attggtccg gaaacatggt tgccctgaa attaacctac aggccaacct ctggcgcggc 18000  
ttctgtatt ccaatattgg gttgtatctt ccagactctc tgaaaatcac ccccgacaac 18060  
atcagctgc cagaaaaca aaacacttat cagtacatga acggtcgcgt aacgccacc 18120

---

-continued

---

gggctcatag acacctatgt aaacgtgggc gcgcgctggt ccccgatgt catggacagc 18180  
attaaccctt tcaaccacca ccgtaacgcg ggcttgcgct accgctccat gctcttgggc 18240  
aacggccggt atgtgccttt tcacattcag gtgccccaaa aattctttgc cattaanaaac 18300  
ctgctgcttc tccccggttc ctatacctat gagtggaaact tccgcaagga tgtcaacatg 18360  
atcctgcaga gctcgcctgg taatgacctg cgagtggacg gggccagcat acgctttgac 18420  
agcattaacc tgtatgcca cttttttccc atggcccaca acacggcctc taccctgga 18480  
gccatgctgc gcaacgacac caatgaccag tccttcaacg actacctgtg cgcggctaac 18540  
atgctgtacc ccatccccgc caacgccacc agcgtgccc tttctattcc tctcgggaa 18600  
tgggctgcct tcaggggctg gagttttact cgcctcaaaa ccaaggagac tccctcgtg 18660  
ggctccgggt ttgaccoccta ctttgtttac tccggctcca tccctacct agatggcacc 18720  
ttttacctca accacacttt caaaaagtg tctattatgt ttgactcctc ggttagctgg 18780  
cccggcaacg accgctgct aacgccaac gagtcgaaa ttaagcgttc cgtggacggt 18840  
gaagggtaca acgtggccca gagcaacatg accaaggact ggtttcta tcaaatgctc 18900  
agtcactata atataggtta ccagggttc tatgtgccc agaactaca ggaccgcatg 18960  
tactccttct tccgcaactt ccaaccaatg agccggcagg tggtagatac cgtgacttat 19020  
acagactaca aagatgtcaa gctcccctac caacacaaca actcagggtt cgtgggctac 19080  
atgggaccca ccatgagaga gggacaggcc taccgggcca actatcccta cccctgatc 19140  
ggagagactg ccgtaccag cctcacgcag aaaaagtcc tctgcgaccg ggtgatgtgg 19200  
aggataccct tctctagcaa ctttatgtcg atgggctccc tcaccgacct ggggcagaac 19260  
atgctgtacg ccaactccgc tcacgccttg gacatgactt ttgaggtgga tcccatggat 19320  
gagcccacgc ttctctatgt tctgtttgaa gtcttcgacg tgggtgcgcat ccaccagccg 19380  
caccgcggcg tcatcgaggc cgtctacctg cgcacacctt tctctgcccg taacgccacc 19440  
acctaagaa gctgatgggt tccagogaac aggagttgca ggccattgtt cgcgacctgg 19500  
gctgcgggccc ctgctttttg ggcaccttcg acaagcgttt tcccgattc atgtcccccc 19560  
acaagccggc ctgcccctc gttaaacacg ccggacggga gacagggggg gtgcactggc 19620  
tcgctctcgc ctggaacccc cgcacccgca cctgctacct gttcgacct tttggtttct 19680  
ccgacgaaa gctgaagcag atctaccaat tcgagtacga gggctcctc aagcgcagcg 19740  
ctctggcctc cagccccgac cactgcgtca cctggaaaa gtccaccag acggtccagg 19800  
ggccctctc ggccgcctgc gggcttttct gttgcatgtt tttgcacgc tctgtgact 19860  
ggcctcacac cccatggag cgcaacccca ccatggatct gctcaccgga gtgccaaca 19920  
gcatgcttca cagtccccag gtcgccccca cctgcgtcg caatcaggac cacctgtatc 19980  
gctttctggg gaaacactct gcctatttcc gccgccaccg gcagcgcac gaaacggcca 20040  
cggccttcca aagcatgagc caaagagtgt aatcaataaa aaccgttttt atttgacatg 20100  
atacgcgctt ctggcgtttt tattaanaat cgaaggttc gagggagggg tctcgtgcc 20160  
cgctggggag ggacacgttg cggtactgga atcgggcgct ccaacgaaac tcggggatca 20220  
ccagccgcgg cagggccacg tcttccatgt tctgcttcca aaactgtcgc accagctgca 20280  
gggtcccat cactgcgggc gctgagatct tgaagtcgca gttagggccg gagccccgc 20340  
ggctgttgcg gaaacagggg ttggcacact ggaacaccaa cacgctgggg ttgtggatac 20400

---

-continued

---

tagccagggc cgtegggtcg gtcacctcag atgcatccag atcctcggca ttgctcaggg 20460  
cgaacggggg cagcttgcaac atctgccgcc cgatctgggg taccaggtcg cgcttgttga 20520  
ggcagtcgca gcgcagaggg atgaggatgc gacgctgccc gcgttgcatg atgggtaac 20580  
tcgcccagga gaactcctct atctgacgga aggccatctg ggccctgacg ccctcgggtga 20640  
aaaaatagccc acaggacttg ctggaaaaca cgttattgcc acagttagatg tcttccgcgc 20700  
agcagcgcgc atcttcgttc ttcagctgaa ccacgttgcg accccagcgg ttctgaacca 20760  
ccttgctttt cgtgggatgc tccctcagcg cccgctgtcc gttctcgtcg gtcacatcca 20820  
ttccaccacac gtgctccttg cagaccatct ccaactccgtg gaaacagaac agaatgccct 20880  
cctggtgggt attgcatgac tcccacacgg cgcacccggg ggactcccag ctcttgtgtt 20940  
tcacccccgc gtaggcttcc atgtaagcca ttagaatct gccatcagc tcagtgaagg 21000  
tcttctggtt ggtgaaggtt agcggcaggg cgcggtgttc ctcgttcaac caagtttgac 21060  
agatcttgcg gtacacggct ccctggctcg gcagaaactt aaaagtcgtt ctgctctcgt 21120  
tgtcacagtg gaacttctcc atcaacatcg tcatgacttc catgcccttc tcccaggcag 21180  
tcaccagcgg cgcgctctcg gggttcttca ccaacacggc ggtggagggg ccctcgcggg 21240  
ccccgacgtc ctcatgagc atttttttaa actccacggg gccgtccgcg cggcgctactc 21300  
tgcgcatcgg agggtagctg aagcccacct ccatgacggg gctttcggcc tcgctgtcgg 21360  
agacgatctc cggggagggc ggcggaacgg gggcagactt gcgagccttc ttcttggggg 21420  
ggagcggagg caactcctgc tcgctcctcg gactcatctc ccgcaagtag ggggtgatgg 21480  
agcttctcgg ttggttctga cggttggcca ttgtatccta ggcagaaaga catggagctt 21540  
atgcccaggg aaactttaac cgccccgtcc cccgtcagcg acgaaaggt catcgtcga 21600  
caggacccgg gctacgttac gccgcccag gatctggagg gcccttaga cgaccggcgc 21660  
gacgctagtg agcggcagga aatgagaaa gaggagggg agggctgcta cctcctggaa 21720  
ggcgacggtt tgctaaagca tttcggcagg cagagcacca tactcaagga ggccttga 21780  
gaccgctccg aggtgcctt ggacgtcgc gcgctctccc aggcctacga ggcgaacctt 21840  
ttctcggccc gactgcctcc gaagagacag cccaacggca cctgcgagcc caaccgcga 21900  
ctcaacttct acccgtgtt cgccgtgccc gaggcgctgg ccacctacca catctttttc 21960  
aaaaaccagc gattccctt ttcctgcccg gccaacggca ccgcgccga taggaagcta 22020  
acaactcagaa acggagtcag catacctgat atcacgtcac tggaggaagt gcctaagatc 22080  
ttcaggggtc tgggtcgaga tgagaagcgg gcggcgaacg ctctgcagaa agaacagaaa 22140  
gagagtcaga acgtgctggt ggagctggag ggggacaacg cgcgtctgac cgtcctcaaa 22200  
cgttgcatag aagtttccca ctctgcctac ccggccctca acctgccgcc caaagttatg 22260  
aaatcgggtc tggaccagct actcatcaag agagctgagc ccctgaatcc cgaccacct 22320  
gaggcggaaa actcagagga cggaaagccc gtcgtcagcg acgaggagct cgagcgggtg 22380  
ctggaacca gggaccccca gcagttgcaa gagaggcga agatgatgat ggcggccgtg 22440  
ctggtcagcg tggagctaga atgctgcaa cggtttttca gcgacgtgga gacgctacgc 22500  
aaaaatcggg agtccctgca ctacacctc cgccagggt acgttcgcca ggcctgcaaa 22560  
atctccaacg tagagctcag caacctggtt tctacatgg gcatcctcca cgagaaccgg 22620  
ctggggcaga gcgtgctgca ctgcaacttg caaggcaggg cgcgaaggga ctacgtccga 22680

---

-continued

---

gactgcgtct acctcttcct caccctcacc tggcagaccg ccatgggcgt gtggcagcag 22740  
tgcttgggaag agagaaacct caaagagctg gacaaactcc tctgccgccca gcggcgggcc 22800  
ctctggaccg gcttcagcga gcgcacggtc gcctgcgccc tggcagacat cattttccca 22860  
gaacgcctga tgaaaacctt gcagaacggc ctgccggatt tcatcagtcga gagcatcttg 22920  
caaaacttcc gctccttcgt cctggagcgc tccgggatct tgcccgccat gagctgcgcg 22980  
ctgccttctg actttgtccc cctttcctac cgcgagtgcc ctccccact gtggagccac 23040  
tgctactctt tccaactggc caactttctg gcctaccact ccgacctcat ggaagacgtg 23100  
agcggagagg ggctgctcga gtgccactgc cgctgcaacc tctgcacccc ccacagatcg 23160  
ctggcctgca acaccgagct gctcagcga acccaggtca taggtacctt cgagatccag 23220  
gggccccagc agcaagaggg tgcttccggc ttgaagctca ctccggcgct gtggacctcg 23280  
gcttacttac gaaaatttgt agccgaggac taccacgccc acaaaattca gttttacgaa 23340  
gaccaatctc gaccaccgaa agccccctc acggcctgcg tcatcaccca gagcaaaatc 23400  
ctggcccaat tgcaatccat caaccaagcg cgccgagatt tccttttgaa aaagggtcgg 23460  
gggggtgacc tggaccccc gaccggcgag gaactcaacc cgtccacact ttccgctcga 23520  
gcagccccc cgagacatgc cacccaaggg aaccgccaag cagctgatcg ctccggcagag 23580  
agcgaagaag caagagctgc tccagcagca ggtggaggac gaggaagagc tgtgggacag 23640  
ccaggcagag gaggtgtcag aggacgagga ggagatggaa agctgggaca gcctagacga 23700  
ggaggacgag ctttcagagg aagaggcgac cgaagaaaa ccacctgcat ccagcgcgcc 23760  
ttctctgagc cgacagccga agccccggc cccgacgccc ccggccggct cactcaaagc 23820  
cagccgtagg tgggacgcc cgggatctcc agcggcagcg gcaacggcag cgggtaaggc 23880  
caaacgcgag cgcggggggt attgtcctg gcggaccac aaaagcagta tctgtaactg 23940  
cttgcaacac tgcgggggaa acatctcctt tgcccgacgc tacctcctct tccatcacgg 24000  
tgtggccttc cctcgcacag ttctctatta ttaccgtcat ctctacagcc cctacgaaac 24060  
gctcggagaa aaaagctaag gcctcctctg ccgagaggaa aaactccgcc gccgctgccg 24120  
ccaaggatcc gccggccacc gaggagctga gaaagcgcct ctttcccact ctgtatgcta 24180  
tctttcagca aagccggggg cagcacctc agcgcgaact gaaaataaaa aaccgctcct 24240  
tccgctcact caccgcgagc tgtctgtacc acaagagaga agaccagctg cagcgcaccc 24300  
tggacgacgc cgaagcactg ttcagcaaat actgctcagc gtctcttaa gactaaaaga 24360  
cccgcgcttt ttccccctcg ggcgcocaaa cccacgtcat cgcagcatg agcaaggaga 24420  
ttcccacccc ttacatgtgg agctatcagc cccagatggg cctggccgcg ggggcccgcc 24480  
aggactactc cagcaaaatg aactggctca gcgcggccc ccacatgatc tcacgagtta 24540  
acggcatccg agcccaccga aaccagatcc tottagaaca ggcggcaatc accgccacac 24600  
cccggcgcca actcaaccgg cccagttggc ccgcccocca ggtgtatcag gaaactcccc 24660  
gcccgaccac agtcctcctg ccacgcgacg cggaggccga agtcctcatg actaactctg 24720  
gggtacaatt agcggggggg tccaggtacg ccaggtacag aggtcgggcc gctccttact 24780  
ctcccgggag tataaagagg gtgatcattc gaggcggagg tatccagctc aacgacgagg 24840  
cggtagactc ctcaaccggc ctccagactg acggagtett ccagctcggg gagcggggcc 24900  
gctcttctct caccactcgc caggcctacc tgaccctgca gagctcttcc tcgcagccgc 24960

---

-continued

---

gctccggggg aatcggcact ctccagttcg tggaagagtt cgtcccctcc gtctacttca 25020  
accogttttc cggctcacct ggacgctacc cggacgcctt cattcccaac tttgacgcag 25080  
tgagtgaatc cgtggacggc tacgactgat gacagatggt gcggccgtga gagctcggct 25140  
gcgacatctg catcactgcc gccagcctcg ctgctacgct cgggagggca tegtgttcag 25200  
ctactttgag ctgccggagc agcaccctca gggaccggct cacgggttga aactcgagat 25260  
tgagaacgcg cttgagtctc acctcatcga cgccttcacc gcccggcctc tcctggtaga 25320  
aaccgaacgc gggatcacta ccataccctt gttctgcacg tgccccacgc cgggattaca 25380  
tgaagatctg tgttgcctc tttgcgctca gtttaataaa aactgaactt tttgccgtac 25440  
cttcaacgcc acgctgtgtt tctccttggt aaaaaacccc aggagtccctt aacttacaca 25500  
tagcaaaacc cttgtattht accatagaaa aacaactagc cctttcaatt ggaaaagggt 25560  
taacaattht tgctacagga cagttgaaa gcacagcaag cgtacaggac agcgcctacac 25620  
caccctacg tggatthtcc cctttaaagc tgacagacaa cggtttaaca ttaagctatt 25680  
cagatcccct gcgtgtggtg ggtgaccaac ttacgtttaa ttttacttct cactacggtt 25740  
acgaaaatgg cagtcttaca ttcaactaca cttctcccat gactactaata aacaacagtc 25800  
ttgctattaa cgtcaatacc tccaaaggcc tcagtagtga caacggcaca ctgcctgtaa 25860  
atgttactcc agatthttaga tthaacagct ctggtgcctt aactthtggc atacaaagtc 25920  
tatggacttht tccaacacaa actcctaact gtaccgtgtt taccgaaagt gactccctgc 25980  
tgagtcttht cttgactaaa tgcggagctc acgtacttgg aagcgtgagt ttaagcggag 26040  
tggcaggaac catgctaaaa atgaccacaa cttctgttac cgttcagtht tegtthtgatg 26100  
acagtggtaa actaatattht tctccacttg cgaacaacac ttgggtgtgt cgacaaagcg 26160  
agagtccgtht gcccaaccca tcttcaacg ctctcacgtht tatgcaaac agtaccattht 26220  
attctagagg agcaagtaac gaacctcaaa acaattatta tgtccagacg tatcttagag 26280  
gcaacgtgag aaagccaatt ctactaactg ttacctacaa ctcagthaat tcaggatatt 26340  
ccttaacttht taaatgggat gctgtcgcca atgaaaaatt tgccactcct acatcttctg 26400  
tttgcctatg tgcagagcaa taaaaccctg ttaccccacc gtctcgttht tttcagatga 26460  
aacgagcggag agttgatgaa gacttcaacc cagtgtaacc ttatgacccc ccatacgtc 26520  
ccgtcatgcc cttcattact ccgccttht cctcctcgga tgggttgag gaaaaaccac 26580  
ttggagtht aagthttaaac tacagggatc ccattactac acaaatggg tctctcacg 26640  
taaaactag agaacgcctc actctaaaca accagggaca gttaacatca actgctggcg 26700  
aagtggagcc tccgctcact aatgctaaca acaacttgc actagcctat agcgaaccat 26760  
tagcagtaaa aagcaaccgc ctaactctat cacacaccgc tccccttgc atcgtaata 26820  
attctthtag gttgcaagtht tcagagccta tthtthttaa tgacgatgac aagctagccc 26880  
tgacagacgc cgcctccctt gtaaccaacg ctggcaccct tcgcttacag agcgcctgccc 26940  
ctthtaggatt ggttgaaaat actcttaaac tgctgttht taaaccctg tathtthtcaa 27000  
atgathttht tgcattagcc attgaaagcc cctggctgt agcagccgca ggtactctga 27060  
ccctacaact tactcctcca ttaaagacta acgatgacgg gctaacacta tccacagtcg 27120  
agccattht tgtaaaaaac ggaaacctag gcttgcaaat atcgcgcctt ttagthttht 27180  
aaaaaacag cctthtctgctt gctattaccc ccccgctgag thtthttht aagcagcccg 27240



---

-continued

---

ttcttggttt gggcttcact tttcccctag ctgtcacaaa caacctcctc tccttaaaca 27300  
tgggagacgg agttaaactt acctataata aactaacagc caatttgggt agggatttac 27360  
aatttgaana cggtgcgatt gccgtaacgc ttactgccga attacctttg caatacacta 27420  
acaaacttca actgaatatt ggagctggcc ttcgttacia tggagccagc agaaaactag 27480  
atgtaaacat taacaaaaat aaaggcttaa cttgggacaa cgatgcagtt attcccaaac 27540  
taggatcggg ctacaattt gaccctaag gcaacatcgc tgttatccct gaaaccgtga 27600  
agccgcaaac gttatggagc actgcagatc cctcgcctaa ctgctcagtg taccaggact 27660  
tggatgccag gctgtggctc gctcttgta aaagtggcga catggtgcat ggaagcattg 27720  
ccctaaaagc cctaaaaggg acggttgctaa atcctacagc cagctacatt tccattgtga 27780  
tatattttta cagcaacgga gtcaggcgta ccaactatcc aacgttgac aacgaaggca 27840  
ccttagctaa cagcgcact tggggatacc gacaggggca atctgctaac actaatgtga 27900  
ccaatgccac tgaatttatg cccagctcaa gcaggtacc cgtgaataaa ggagacaaca 27960  
ttcaaaatca atctttttca tacacctgta ttaaaggaga ttttgctatg cctgtcccgt 28020  
tccgtgtaac atataatcac gccctggaag ggtattccct taagttcacc tggcgcgttg 28080  
tagccaatca ggctttgat attccttgct gttcattttc atacatcaca gaataaaaaa 28140  
ccactttttc attttaattt ctttttattt tacacgaaca gtgagacttc ctccaccctt 28200  
ccatttgaca gcatacaca gcctctcccc ctctatagca gtaaactggt gtgaatcagt 28260  
ccggtatttg ggagttaaaa tccaacagc ctctttggtg atgaaacgtc gatcagtaat 28320  
ggacacaaat ccttgggaca ggttttcaa cgtttcgtg aaaaactgca caccgcctta 28380  
caaaacaaac aggttcagc tctccacggg ttatctcccc gatcaaacct agacagggta 28440  
aaggtgcggt ggtgttccac taaaccacgc aggtggcgct gtctgaacct ctcggtgcga 28500  
ctcctgtgag gctggtgaaga agttagattg tccagtagcc tcacagcatg tatcatcagt 28560  
ctacagatgc gtctggcgca gcagcgcac tgaatctcac tgagattccg gcaagaatcg 28620  
cacaccatca caatcaggtt gttcatgatc ccatagctga acacgctcca gccaaaagctc 28680  
attcgtctca acagcggcc cgcgtgtccg tccaacctta ctttaacata aatcaggtgt 28740  
ctgcccgtga caaacatgct acccacatag agaacttccc ggggcaggcc cctgttcacc 28800  
acctgtctgt accagggaaa cctcacattt atcagggagc catagatggc ctttttaaac 28860  
caattagcta ataccgccc accagctcta cactgaagag aaccgggaga gttacaatga 28920  
cagtgaataa tccatctctc ataaccctg atggtctgat gaaaatctag atctaactg 28980  
gcacaacaaa tacacacttt catatacatt ttcataacat gtttttccca ggccgttaaa 29040  
atacaatccc aatacagggg ccaactcctgc agtacaataa agctaataca agatggtata 29100  
ctcctcacct cactgacact gtgcagtctc atattttcac attctaagta ccgagagttc 29160  
tcctctacag cagcactgct gcggtcctca caaggtggta gctggtgatg attgtagggg 29220  
gccagctcgc agcgataccg tctgtcgcgt tgcacgttag accaggaacc gacgcacctc 29280  
ctcgtacttg tggtagcaga accacgtccg ctgccagcac gtctccaagt aacgccggtc 29340  
cctcgtcgc tcacgtctcc tctcaatgc aaagtcaac cactcttgta atccacacag 29400  
atccctctcg gcctccgggg tgatgcacac ctcaaaceta cagatgtctc ggtacagttc 29460  
caaacacgta gtgagggcga gttccaacca agacagacag cctgatctat cccgacacac 29520

-continued

---

```

tggagggtgga ggaagacacg gaagaggcat gttattccaa gcgattcacc aacgggtcga 29580
aatgaagatc ccgaagatga caacgggtcgc ctccggagcc ctgatggaat ttaacagcca 29640
gatcaaacgt tatgcgatcc tccaagctat cgatcgccgc ttccaaaaga gcctggaccc 29700
gcacttccac aaacaccacg aaagcaaaag cactattatc aaactcttca atcatcaagc 29760
tgcaggactg tacaatgcct aagtaatfff cgtttctcca ctccggaatg atgtcggcgc 29820
agatagtctg aaggttcatc ccgtgcaggg taaaagctc cgaagggcg ccctctacag 29880
ccatgcgtag acacaccatc atgactgcaa gatatcgggc tcctgagaca cctgcagcag 29940
atttaacaga tcaaggtcag gttgtctccc gcgatcacga atctccatcc gcaaggtcat 30000
ttgcaaaaaa taaataaat ctatgccac tagatctgtc aactccgat taggaaccaa 30060
atcagggtgtg gctacgcagc acaaaagttc cagggatggt gccaaactca ctagaaccgc 30120
tcccagagtaa caaaactgat gaatgggagt aacacagtgt aaaatgtgca accaaaaatc 30180
actaagggtg tcctttaaaa agtccagtac ttctatattc agtccgtgca agtactgaag 30240
caactgtgcg ggaatatgca caacaaaaa aatagggcgg ctccagatca tgttgacctc 30300
aaataaaaaa aatcattaaa ctaaagaagc ttggcgaacg gtgggataaa tgacacgctc 30360
cagcagcaga caggcaaccg gctgtccccg ggaaccgcgg taaaattcat ccgaatgatt 30420
aaaaagaaca acagaaactt cccaccatgt actcggttgg atctcctgag cacacagcaa 30480
tccccccctc acattcatgt ccgccacaga aaaaaacgt cccagatacc cagcggggat 30540
atccaacgac agctgcaaag acagcaaac aatccctctg ggagcgatca caaaatcctc 30600
cggtgaaaaa agcacataca tattagaata accctggtgc tggggcaaaa aggcccgcg 30660
tcccagcaaa tgcacataaa tatgttcatc agccattgcc ccgtcttacc gcgtaatcag 30720
ccacgaaaaa atcgagctaa aattcaccca acagcctata gctatatata cactccgccc 30780
aatgacgcta ataccgcacc acccagcacc aaagttcacc cacaccaca aaaccgcga 30840
aaatccagcg ccgtcagcac ttccgcaatt tcagtctcac aacgtcaact ccgcgcgct 30900
tttcacattc ccacacacac ccgcgccctt cgccccgcc tcgcgccacc ccgcgtcacc 30960
gcacgtcacc ccggccccgc ctgcctcctc cccgctcatt atcatattgg cacgtttcca 31020
gaataaggta tattattgat gatg                                     31044

```

```

<210> SEQ ID NO 6
<211> LENGTH: 34115
<212> TYPE: DNA
<213> ORGANISM: simian adenovirus SV-39

```

```

<400> SEQUENCE: 6

```

```

catcatcaat ataacaccgc aagatggcga ccgagttaac atgcaaatga ggtgggcgga 60
gttacgcgac ctttgtcttg ggaacgcgga agtgggcgcg gcgggtttcg gggaggagcg 120
cggggcgggg cgggcgtgtc gcgcggcggg gacgcgccgg ggaccggaa attgagtagt 180
ttttattcat ttgcaagtt tttctgtaca ttttggcgcg aaaactgaaa cgaggaagtg 240
aaaagtgaaa aatgccgagg tagtcaccgg gtggagatct gacctttgcc gtgtggagtt 300
taccgctga cgtgtggggt tcggtctcta ttttttact gtggttttcc ggttacggtc 360
aaagttccc attttatgac tccacgtcag ctgatcgeta gggatattaa tgcgcctcag 420
accgtcaaga ggccactctt gactgcccgc gagaagagtt ttctcctccg cgttccgcca 480

```

-continued

---

|  |      |
|--|------|
| actgtgaaaa aatgaggaac ttcttgctat ctccggggct gccagcgacc gtagccgccg  | 540  |
| agctgttgga ggacattgtt accggagctc tgggagacga tcctcaggtg atttctcact  | 600  |
| tttgtgaaga ttttagtctt catgatctct atgatattga tccgggtggt gaggggcaag  | 660  |
| aggatgaatg gctggagtct gtggatgggt tttttccgga cgctatgctg cttagaggctg | 720  |
| atttgccacc acctcacaac tctcacactg agcccagatc agctgctatt cctgaattgt  | 780  |
| catcaggtga acttgacttg gcttgttacg agactatgcc tccggagtcg gatgaggagg  | 840  |
| acagcgggat cagcgatccc acggctttta tggctcttaa ggcgattgct atactaaaag  | 900  |
| aagatgatga tggcgatgat ggatttcgac tggacgctcc ggcggtgccg gggagagact  | 960  |
| gtaagtccctg tgaataccac cgggatcgtg ccggagaccc gtctatgctg tgttctctgt | 1020 |
| gttatctccg tcttaacgct gcttttgtct acagtaagtg ttttgtgctt ttttaccctg  | 1080 |
| tggctttggt gaggtttatt ttttctgtgt ctcatagggg gttgtttatt ataggtcctg  | 1140 |
| tttcagatgt ggaggaacct gatagtacta ctggaaatga ggaggaaaag ccctccccgc  | 1200 |
| cgaactaac tcagcctgc agacctaata ttttgagacc ctccggcccag cgtgtgtcat   | 1260 |
| cccggaaaac tgctgctggt aattgcatag aagatttatt ggaagagccc actgaacctt  | 1320 |
| tggacttgct cttaaagcga ccccgcccg agtagggcgc ggtgccagtt ttttctctct   | 1380 |
| agcttccggg tgactcagtg caataaaaat tttcttgca acaggtgtat gtgtttactt   | 1440 |
| tacggggcgg aagggattag gggagtataa agctggaggg gaaaaatctg aggctgtcag  | 1500 |
| atcgagtgag aagtccatg gacttgtagc agagcctaga gaatctaagt tctttgcgac   | 1560 |
| gtttgctgga ggaggcctc gacagaacct cttacatttg gaggtttctg ttcggttccc   | 1620 |
| ctctgagtcg ctttttgca cgggtgaagc gagagcacct gacggaattt gatgggcttt   | 1680 |
| tagagcagct gcctggactg tttgattctt tgaatctcgg ccaccggacg ctgctagagg  | 1740 |
| agaggctttt tccacaattg gacttttctc ctccaggcgg tctgtgttca gcgcttgctt  | 1800 |
| ttgctgtaca tctgttgac agatggaacg agcagacgca gctcagccc ggttactc      | 1860 |
| tggacttcct gacgctatgc ctatggaagt tcggaatcag gaggggagg aagctgtacg   | 1920 |
| ggcgcttggt ggagaggcat ccgtctctgc gccagcagcg tctgcaagct caagtgctgc  | 1980 |
| tgaggcggga gatctgtaa gccatttcgg aggagagag cggcatgga gagaagaatc     | 2040 |
| cgagagcggg gctggacct ccggcggagg agtagggggg ataccggacc cttttcctga   | 2100 |
| gttgctttg gggcggtg ggggccttc tgtgtacgt gaggatgaag agggcgcca        | 2160 |
| acgggtcag aagagggagc attttgatc ctgcacttc ttgctgatg taaccgtggc      | 2220 |
| cctgatggcg aaaaacagc tggaggtggt gtgtaccgg gaagtatggg aggactttga    | 2280 |
| gaagggggac ttgcacctgc tggaaaaata taactttgag caggtgaaa cactactggat  | 2340 |
| gaaccggat gaggactgg aggtggtttt gaaccgatac ggcaaggtag ctctgcgtcc    | 2400 |
| cgactgtcgc taccaggttc gcgacaaggt ggtcctcga cgcaactgt acctgtggg     | 2460 |
| caacggcgcc accgtggaga tggtggaacc cagaaggggt ggttttgtg ccaatatgca   | 2520 |
| agaaatgtgc cctgggtg tgggctgtc tgggtgact tttcatagt tgaggtttag       | 2580 |
| cggtagcaat tttgggggtg tggttattac cgcgaacct cctgtgtcc tgcataattg    | 2640 |
| ctactttttt ggcttcagca acacctgtg gaaatgagg gtggaggca aagtgcggg      | 2700 |
| gtgttccctt tacgcttct ggaaggggt ggtgagccag gtaaggcta aagtgtctgt     | 2760 |

-continued

---

|   |      |
|---|------|
| tcacaagtgt atgttggaga gatgcacctt gggcatttcc agtgagggct tcctccacgc | 2820 |
| cagcgacaac gtggcttctg acaacggctg cgcctttctt atcaagggag ggggtcgc   | 2880 |
| ctgtcacaac atgatatgcg gccctgggga tgtccccca aagccttacc agatggttac  | 2940 |
| ctgcacagat ggcaagggtc gcatgctcaa gcctgtgcac attgtgggcc accggcgcca | 3000 |
| ccgctggcca gagtttgaac acaatgtgat gacctgctgt agcttgtacc tgggaggcag | 3060 |
| gcgaggagtt ttcttgccca gacagtgtaa cctggccac tgcaacgtga tcatggaaca  | 3120 |
| atccgcgct acccaggttt gctttggagg aatatttcat ataagcatgg tgggtataa   | 3180 |
| gactctgctg tacgacgact gtcgggctcg tactcgaacc tgcgactgcg gagcctctca | 3240 |
| cctgtgtaac ctgactgtga tgggatggt gactgaggag gtgcgactgg accactgtca  | 3300 |
| gcactcttgc ctgccccgag agttttcttc ctcggacgag gaggactagg taggtggtt  | 3360 |
| ggcggtggcc agcgagaggg tgggtataa aggggaggtg tcgctgacg ctgtcttctg   | 3420 |
| ttttcaggt accatgagcg gatcaagcag ccagaccgcg ctgagcttcg acggggccgt  | 3480 |
| gtacagcccc tttctgacgg ggcgcttgcc tgcctgggcc ggagtgcgtc agaattgtac | 3540 |
| cggttcgacc gtggacggac gtcccgtgga tccatctaac gctgcttcta tgcgctacgc | 3600 |
| tactatcagc acatctactc tggacagcgc cgctgcccgc gcagccgcca cctcagccgc | 3660 |
| tctctcgcgc gccaagatca tggctattaa cccaagcctt tacagccctg tatccgtgga | 3720 |
| cacctcagcc ctggagcttt accggcgaga tctagctcaa gtggtggacc aactcgcagc | 3780 |
| cgtagccaa cagttgcagc tgggtgcgac ccgagtgag caactttccc gccctccca    | 3840 |
| gtaaccgcaa aaattcaata aacagaatth aataaacagc acttgagaaa agtttaact  | 3900 |
| tgtggtgac tttattcctg gatagctggg gggagggaa ggcgggaac gtaagacctg    | 3960 |
| gtccatcggt cccggtcggt gagaacacgg tggatthttt ccaagaccg atagaggtg   | 4020 |
| gtctgaacgt tgagatacat gggcatgagc ccgtctcggg ggtggaggta ggcactg    | 4080 |
| agggcctcgt tttcaggggt ggtgtgtaa atgatccagt cgtaggcccc ccgctggcg   | 4140 |
| tgggtcgtga agatgcctt cagcagcaag ctgatggcaa cgggaagacc cttggtgtag  | 4200 |
| gtgttgaca agcggttgag ttgggagggg tgcagcggg gactgatgag gtgcatttt    | 4260 |
| gcctggatct tgaggttggc tatgttccg cccagatcgc gcctgggatt catgttatgc  | 4320 |
| aagaccacca gcaccagta accggtgcag cgggggaatt tgcctgcag cttggaagg    | 4380 |
| aaagcgtgga agaatttggg gaccctcgg tgcctgccta ggttttccat gactcatcc   | 4440 |
| atgatgatgg cgatgggccc ccgggaggca gcctgggcaa aaacgttgcg ggggtccgtg | 4500 |
| acatcgtagt tgtggtcctg ggtgagttca tcataggaca tttgacaaa gcgcggcag   | 4560 |
| aggtcccg actggggaat gatggttcca tccggtccgg gggcgtgatt gccctcgag    | 4620 |
| atgtgcatth cccaggttth gatthcagag ggaggatca tgtcaacctg gggggcgatg  | 4680 |
| aaaaaatgg tctctggggc ggggtgatg agctgggtgg aaagcaggtt gcgcaagagc   | 4740 |
| tgtgacttgc cgcagccggt gggcccgtag atgacageta tgacgggttg cagggtgtag | 4800 |
| tttagagagc tacaactgcc atcatccttc aaaagcgggg ccacactgtt taaaagttct | 4860 |
| ctaactgta agttttccc cactaagtc tgcaggagac gtgacctcc tagggagaga     | 4920 |
| agttcaggaa gcgaagcaaa gtttttaagt ggcttgaggc catcgccaa gggcaagttc  | 4980 |
| ctgagagttt gactgagcag ttccagccgg tcccagagct cggttacgtg ctctacggca | 5040 |

---

-continued

---

tctcgatcca gcagacctcc tcgtttcggg gggtggggcg gctctggctg tagggaatga 5100  
ggcgggtggc gtccagctgg gccatggtgc ggtccctcca tgggcgcagg gttctcttca 5160  
gggtggtctc ggtcacggtg aatgggtggg ccccgggctg ggcgctggcc agggtgcgct 5220  
tgaggctgag gcggctggtg gcgaaccgtt gcttttcgtc tccctgcaag tcagccaaat 5280  
agcaacggac catgagctca tagtccaggc tctctgcggc atgtcctttg gcgcgaagct 5340  
tgcctttgga aacgtgcccg cagtttgagc agagcaagca ttttagcgcg tagagttttg 5400  
gcgccaagaa cacggtattcc ggggaataag catccccacc gcagttggag caaacggttt 5460  
cgcatccacc cagccaggtc agctgaggat cttttgggtc aaaaaccaag cgcccgccgt 5520  
ttttttgat gcgcttccta cctcgggtct ccatgaggcg gtgcccgcgt tcggtgacga 5580  
agaggctgtc ggtgtctccg tagacggagg tcaggggcgc ctcctccagg ggggtcccgc 5640  
ggtcctcggc gtagagaaac tcgcaccact ctgacataaa cgcccgggtc caggctagga 5700  
cgaatgaggc gatgtgggaa gggtaccggt cgttatcgat gagggggtcg gttttttcca 5760  
aggtgtgcag gcacatgtcc cctcgtccg cttccaaaaa tgtgattggc ttgtaggtgt 5820  
aagtcaagtg atcctgtcct tcccgggggg tataaaaggg ggcgtttccc cctcctcgt 5880  
cactctcttc cggttcgctg tcgccaaggc ccagctgttg ggttacgtaa acgcggtga 5940  
aggcgggcat gacctgtgcg ctgaggttgt cagtttctat atacgaggaa gatttgatgg 6000  
cgagcggccc cgtggagatg ccctgagggt gctcggggcc catttggtca gaaaacacaa 6060  
tctgtcggtt atcaagcttg gtggcaaaag acccgtagag ggcgttgag agcaacttg 6120  
cgatggagcg ctgggttttg tttttttccc ggtcggcttt ttccttgcc gcgatgttga 6180  
gctggacgta ctccctggcc acgcacttcc agccgggaaa aacggccgtg cgctcgtccg 6240  
gcaccagcct cacgctccat ccgcggttgt gcagggtgat gacgtcgatg ctggtggcca 6300  
cctctcgcg caggggctcg ttggtccagc agaggcgacc gcccttgcga gacgagaag 6360  
ggggcagggg gtcaagcagg cgctcgtccg gggggtcggc gtcgatggta aagatggcgg 6420  
gcagcaggty tttgtcaaa taatcgatct gatgcccggg gcaacgcagg gcggtttccc 6480  
agtcccgcac cgccaaggcg cgctcgtatg gactgagggg ggcgcccag ggcagggat 6540  
gcgtcagggc cgaggcgtac atgccgcaga tgtcatagac gtaaaggggc tctccagga 6600  
cgccagggta ggtggggtag cagcgcccc cgcgatgct ggcccgtacg tagtcgtaga 6660  
gctcgtgcga gggggccaga aggtggcggc tgaggtgagc gcgctggggc tttcatctc 6720  
ggaagaggat ctgcctgaag atggcgtggg agttggagga gatggtgggc cgctgaaaaa 6780  
tgttgaagcg gcgctcgggc agaccacgg cctcggcgtat aaagtggcg taggactctt 6840  
gcagcttttc caccagggag gcggtgacca gcacgtccag agcgcagtag tccagggttt 6900  
cccgcacgat gtcataatgc tcttcctttt tttccttcca gaggtctcgg ttgaagagat 6960  
actcttcgcy gcttttccag tactcttggg gaggaaccoc gttttcgtct ccacggtaag 7020  
agcccaacat gtaaaactgg ttgacggcct gatagggaca gcaccccttc tccacgggca 7080  
gcgagtaggc cagggcggcc ttgcccaggg aggtgtgagt cagggcaaaag gtgtcgcgga 7140  
ccataacttt tacaaactgg tacttaaaat cccggtcgtc gcacatgcct cgctcccagt 7200  
ctgagtagtc tgtgcgcttt ttgtgcttgg ggttaggcag ggagtaggtg acgtcgttaa 7260  
agaggattht gccacatctg gccataaaat tgcgagagat tctgaagggg ccgggcacct 7320

---

-continued

---

ccgagcgggtt gttgatgact tgggcagcca ggagaatttc gtcgaagccg ttgatgttgt 7380  
gccccacgac gtagaactct atgaaacgcg gagcgccgcg cagcaggggg cacttttcaa 7440  
gttgctggaag agtaagtcc cgcggtcga cgcggttcc cgtgcggtc cagtcctcca 7500  
ccgggtttcg ctcccacaaa tcctgccaga tgtggtcgac tagcaagagc tgcagtcggt 7560  
cgcgaaattc gcggaatctt ctgccgatgg cttgcttctg ggggttcaag caaaaaaagg 7620  
tgtctcgtg gtgcgcacc gctgccagc cgagctcgcg agccagattc agggccagca 7680  
gcaccagagc cggctcaccg gtgattttca tgacgaggag aaagggcacc agctgttttc 7740  
cgaacgcgcc catccagggt taggtctcca cgtcgtaggt gagaacaga cgttcgggtc 7800  
gcgggtgcca tcccaggggg aaaaacttga tgggctgcca ccattgggag ctctgggctg 7860  
ggatgtgatg gaagtaaaag tcccggcgcc gcgtggaaca ttcgtgctgg tttttgtaa 7920  
agcggccgca gtggtcgcag cgcgagacgg agtgaaggct gtgaatcagg tgaatcttgc 7980  
gtcgtcgtg gggccccaga gccaaaaagc ggagcgggaa cgaccgcgcg gccacttcg 8040  
cgtccgagc caagatggat gagggttcca ccgttcccgc cccgcggacc gaccagactt 8100  
ccgccagctg cggcttcagt tcttgacca gctctcgcag cgttctcgtc ctgggcgat 8160  
cgtgaatacg gaagtgtcgc ggtagagcgc ggagcgggtg gacttccagg aggtgtgtga 8220  
ggcccgagc gagatgcagg tggacttga tttcccacgg atgacggtcg cgggcgtcca 8280  
aggcgaagag atgaccgtgg gcccgcggcg ccaccagcgt tccgcggggg gtctttatcg 8340  
gcggcgggga cgggctcccg gcggcagcgc cggctcggga cccgcgggca agtcgggag 8400  
cggcacgtcg gcgtggagct cgggcagggg ctggtgctgc gcgcggagct gactggcaaa 8460  
ggctatcacc cggcgattga cgtcctggat ccggcggcgc tgcgtgaaga ccaccggacc 8520  
cgtggtcttg aacctgaaag agagttcgac agaataatc tcggcatcgt taaccgcggc 8580  
ctggcgagc atttcggcca cgtcccggga gttgtcttga tacgcgattt ctgccatgaa 8640  
ctggtcgatt tcctcttctc gcaagtctcc gtgaccggcg cgttcgacgg tggccgcgag 8700  
atcgttgag atgcccacca tgagctggga aaaggcattg atgccacct cgttccacac 8760  
tcggctgtac accacctctc cgtgaacgtc gcgggcgcgc atcaccacct gggcgagatt 8820  
gagttccacg tggcgggcca aaaccggata gtttcggagg cgctgataca gatagttgag 8880  
ggtggtggcg gcgtgctcgc ccacaaaaaa atacatgatc cagcggcgga gggtcagctc 8940  
gttgatgtcg cccagcgcct ccaggcgttc catggcctcg taaaagtcca cggcaaatgt 9000  
gaaaaattgg ctgttctcgg ccgagaccgt gagctcttct tccaagagcc gaatgagatc 9060  
cgccacgggtg gccctgactt cgcgttcgaa agccccgggt gcctcctcca cctcttctc 9120  
ctcgacttct tcgaccgctt cgggcacctc ctcttctcgc accaccacct caggcggggc 9180  
tcggcggcgc cggcggcgga cgggcaggcg gtcgacgaaa cgctcgatca tttccccct 9240  
ccgtcgacgc atggtctcgc tgacggcgcg accctgttcg cgaggacgca ggggtaaggc 9300  
gcccggcccg agcggaggta acaggagat cggggggcgg tcgtggggga gactgacggc 9360  
gctaactatg catctgatca atgtttcgt agtgacctcg ggtcggagcg agctcagcgc 9420  
ttgaaaatcc acgggatcgc aaaaccgttc caggaacgcg tctagccaat cacagtcgca 9480  
aggtaagctg aggaccgtct cgggggcttg tctgttctgt cttcccgcgg tgggtctgct 9540  
gatgaggtag ttgaagttag cgtctttgag gcggcggatg gtgacagga gaaccacgtc 9600

---

-continued

---

tttgcccca gcttgctgta tccgcaggcg gtcggccatg cccacactt ctccttgaca 9660  
gcgccggagg tccttgtagt attcttgcac cagcctttcc acgggcacct cgtcttcttc 9720  
ttccgctcgg ccggacgaga gccgcgtcag gccgtaccgg cgtgcccctt gtggttgagg 9780  
caggcccagg tcggccacga cgcgctcggc cagcacggcc tgctggatgc gggtaggggt 9840  
gtcctgaaa gctgcgagat ccacaagcg gtggtacgcg ccagtgttga tgggttaggt 9900  
gcagttgctc atgacggacc agtttacggt ctgggtgcca tggcccacgg tttccaggta 9960  
gcgagagcgc gtagtaggcc gcgtctcga gatgtagtgc ttgcaggacc gcagcaggta 10020  
ctggtagccc accagcagat gcggcggcg ctggcggtag aggggccacc gctgggtggc 10080  
gggggcggtg ggggcgagat cttccaacat gaggcggtga tagccgtaga ttagcgcga 10140  
catccaagt atgcccctgg ccgtggtgct ggcgcggcg tagtcgcga cgcggttcca 10200  
gatgtttcg agcggctgga agtactcgt ggtggggcga ctctgcccgg tagggcgggc 10260  
gcagtcggcg atgctctac gggaaaaaga agggccagt aacaaccgcc ttcgtagcc 10320  
ggaggagaac gcaaggggt caaagaccac cgaggctcgg gttcgaacc cgggtggcgg 10380  
cccgaatac gagggcggtt ttttgctttt ttctcagat catccctgc tgcggcagat 10440  
gcgtccgaac gcgggtccc agtccccgc ggtgcctgcg gccgtgacgg cggtctctac 10500  
ggccacgctc cgtccaccac cgcctaccac ggcccaggcg gcggtggctc tgcgcggcgc 10560  
aggggaacc gaagcagagg cgggtttgga cgtggaggag ggcagggggt tggctcggct 10620  
ggggccctg agtcccgagc ggcaccgcg cgtggctctg aagcgcgacg cggcggaggc 10680  
gtacgtgccc cgagcaatc tgtttcgcga ccgacggcg gagggggcg aggagatgcg 10740  
agacttgct tttcggcg ggaggaggt gcgtcacgg ctggaccggc agagggttct 10800  
gagagaggag gactttgagg cggacgagcg cacgggggtg agtcccgcg gggctcacgt 10860  
ggcggccgc aacctggtga gcgctacga gcagacggtc aaggaggaga tgaacttcca 10920  
gaagagctc aatcatcac tgcgcacgct gattgcgcgc gaagaggtg ccatcgccct 10980  
catgcatctg tgggattttg tggaggcgta cgttcagaac cccagcagca agccgctgac 11040  
ggctcagctg ttctctatc tgcaacatag tcgagacaac gaaacgttca gggaggccat 11100  
gctgaacatt gcagacgct agggcgctg gctcttgat ctattaaca tcttgacag 11160  
tatcgtagt caggagcgt cgtgagcct ggccgacaag gtggtgcca tcaactacg 11220  
catgctgctc ctgggcaaat ttacgccc caagatctac aagtctcgt tctccccat 11280  
agacaaggag gtgaagatag acagctttta catgcgcatg gcgctcaagg tgctgactct 11340  
aagcgcgac ctgggggtgt accgcaacga ccgcatacac aaggcgggtg gcgccagccg 11400  
ccggcgcgag ctgagcagacc gcgagctttt gcacagcctg catcgggctg tgaactgctc 11460  
cggcagcgc gaggcggcg agtactttga cgcggagcg gacttgctct ggcagccatc 11520  
ccgacgcgcg ctggaggcgg ctggcgtcgg ggagtacggg gtcgaggacg acgatgaagc 11580  
ggacgacgag ttgggcatg acttgtagcc gttttctggt agatatgctg gcgaacgagc 11640  
cgtctcggc gccatgggt acggcggcg gcgcgccca ggaccggcc acgocgcg 11700  
cgctgcagag tcagccttcc ggagtgacgc ccgaggacga ctggtccgag gccatgctc 11760  
gcatcctgct gctgacggcg gcacaaccgg aggtttctc gcagcagccg caggcaaac 11820  
ggtttcggc cattttgaa gcggtggtgc cctccagacc caaccacc cagaaaagg 11880

---

-continued

---

tgctggccat cgtaacgcc ctggcggaga ccaaggccat ccgccagac gaggccgggc 11940  
aggtttacaa cgcgctgcta gaaaggggtg gacgctacaa cagctccaac gtgcagacca 12000  
atctggaccg ctgtgtgacg gacgtgaagg aggccgtagc ccagcgagag cggtttttca 12060  
aggaagccaa tctgggctcg ctggtggccc tcaacgcctt cctgagcagc ctgccggcga 12120  
acgtgccccg cggtcaggag gactacgtga actttctgag cgcctccgc ctgatggtgg 12180  
ccgaggtgcc gcagagcgag gtgtaccagt ctggcccaa ctactactt cagacctccc 12240  
ggcagggcct gcagagcgta aacctgacgc aggcctttca gaacctgag gccctttggg 12300  
gggtgcgcgc tccgctgggc gaccgcagca cgggtgcccag cctgctgacc cccaatgccc 12360  
ggctgctctt gtttctcatt gctccgttca ccgacagcgg ttccatcagc cgcgactctt 12420  
acctgggaca cctgctcacc ctgtaccggg aggcctcctg gcagggcgcg gtggacgagc 12480  
agacgtacca ggaaatcacc agcgtgagcc gcgcgctggg gcaggaggac acgggcagct 12540  
tgaggcgcag tctgaacttc ctgctgacca accggcggca gcgcctaact cccagtagc 12600  
cgctgaacgc ggaggaggag cgcctcctgc gtttcgtgca gcagagcacc gcgctgtact 12660  
tgatgcggga agggcctctt cccagcgtt cgctggacat gacggcggcc aacatggagc 12720  
catcgttcta cgcgccaaac cgtcccttcg tcaaccgctt aatggactat ttgcatcggg 12780  
cggcggccct gaaccggaa tactttacta acgtcatcct gaacgaccgt tggctgccac 12840  
ctcccgctt ctacacgggg gagttcgacc tcccggaggc caacgacggt ttcattgtgg 12900  
acgacgtgga cagcgtgttc ctgcccgca agaaggaggc ggggtactct cagagccacc 12960  
gcgcgagcct cgcagacctg ggggcgaccg ggcccgcgtc tccgctgcct cgcctgccga 13020  
gcgccagcag ccccagcgtg gggcgggtga gccgtccgcg cctcagcggg gaggaggact 13080  
ggtggaacga tccgctgctc cgtccggccc gcaacaaaaa cttcccaac aacgggatag 13140  
aggatttggt agacaaaatg aaccgttggg agacgtatgc ccaggagcat cgggagtggc 13200  
agggagggca acccatgggc cctgttctgc cgcctctcgc gcgccgcgc agggacgaag 13260  
acgccgacga ttcagccgat gacagcagcg tgttgatctt gggcgggagc gggaaaccct 13320  
ttgccacct gcaacctcgc ggcgtgggtc ggcgtggcg ctaggaaaaa aaattattaa 13380  
aagcacttac cagagccatg gtaagaagag caacaaaggc gtgtcctgct ttcttcccg 13440  
tagcaaaatg cgtcgggcgg tggcagttcc ctccgcggca atggcgttag gcccgcccc 13500  
ttcttacgaa agcgtgatgg cagcggccc cctgcaagcg ccgttgaga atccttacgt 13560  
gcccgccgca taactggagc ctacgggcgg gagaacagc attcgttact cggagctgac 13620  
gcccctgtac gacaccacc gcctgtacct ggtggacaac aagtcagcag atatgccac 13680  
cttgaactac cagaacgacc acagcaactt tctcacgtcc gtggtgcaga acagcgacta 13740  
cacgcccgcc gaagcgagca cgcagaccat taacttgac gaccgctgc gctggggcg 13800  
ggacttgaaa accattctgc aactaactt gcccaacgtg aacgagttca tgtttacaa 13860  
ctcgttcagg gctaaactta tgggtggcga cgaggccgac aaggaccgg tttatgagt 13920  
ggtgcagctg acgctgccg aggggaactt ttcagagatt atgaccatag acctgatgaa 13980  
caacgccatt atcgaccact acctggcggg agccagacag caggggggta aagaaagcga 14040  
gatcggcgtc aagtttgaca cgcgcaactt tctgtgggc tgggaccgg agacggggct 14100  
tgtgatgccg ggggtgtaca cgaacgaagc tttccatccc gacgtggtcc tcttgccggg 14160



---

-continued

---

ctgcggggtg gactttacct acagccggtt aaacaacctg ctaggcatac gcaagagaat 14220  
gccctttcag gaagggtttc agatcctgta cgaggacctg gagggcggtg acatccccgc 14280  
cctgctggac gtgccggcgt acgaggagag catcgccaac gcaagggagg cggcgatcag 14340  
gggcgataat ttcgcggcgc agccccaggc ggctccaacc ataaaaccg ttttgaaga 14400  
ctccaaaggc cggagctaca acgtaatagc caacaccaac aacacggctt acaggagctg 14460  
gtatctggct tataactacg gcgacccgga gaagggggtt agggcctgga ccctgctcac 14520  
cactccggac gtgacgtgcg gttcagagca ggtctactgg tcgctgctg acatgtacgt 14580  
ggacctgtg acgtttcgt ccacgcagca agttagcaac taccagtggt tgggagcggg 14640  
gcttatgccg attcacagca agagctttta caacgagcag gccgtctact cacagctcat 14700  
tcgtcagacc accgcccata cgcacgtttt caaccgcttc cccgagaacc aaatcctagt 14760  
gcgacctcca gcgcccacca tcaccaccgt cagcgagaac gtgcccgtc taaccgatca 14820  
cgggacgctg cctttgcaga acagcatccg cggagttcag cgagttacca tcacggacgc 14880  
ccgtcgtcgg acctgtccct acgtctacaa agccttggga atcgtggccc cgcgcgtcct 14940  
gtcagtgctg actttctaga tgtccatcct catctctccc agcaacaata ccggttgggg 15000  
tctgggctgt accaaaatgt acggaggcgc caaacgacgg tccccacaac atcccgtgcg 15060  
agtgcgctgg cactttagag ccccattggg gtcgcacacg cgcgggcgca cggcccgaac 15120  
cacgctcgac gacgtgatcg atagcgtggt ggccgacgcc cgcaactacc agcccgtcgt 15180  
atccacggtg gacgaagtca tcgacggcgt ggtggccgac gccaggcctt acgcccgcag 15240  
aaagtctcgt ctgcgcccgc gccgttcgct aaagcgcgcc acggccgcca tgaagccgc 15300  
tcgctctctg ctgcgtcgcg cacgtatcgt gggtcgccgc gccccagac gcgcagccgc 15360  
caacgcccgc gccggccgag tgcgcgccgc ggccgcccag caggccgccg ccgccatctc 15420  
cagtctatcc gcccccgac gcgggaatgt gtactgggtc agggactcgg ccaccggcgt 15480  
gcgagttccc gtgagaacct gtcctctcgt tcctgaata aaaagttcta agcccaatcg 15540  
gtgttccggt gtgtgttcag ctcgtcatga ccaaacgcaa gtttaaagag gagctgctgc 15600  
aagcgtggt ccccgaatc tatgcgccg gcgccgacgt gaaaccgct cgcgtgaaac 15660  
gctgaaagaa gcaggaaaag ctagagacaa aagaggaggc ggtggcgtt ggagacgggg 15720  
aggtggagtt tgtgcctcgc ttcgcgccgc gtcgagcagt gaattggaag gggcgcaagg 15780  
tgcaacgggt gctgcgtccc gccacggtgg tgtctttcac cccgggtgaa aaatccgctt 15840  
ggaagggcat aaagcgcgtg tacgatgagg tgtacgggga cgaagacatt ctggagcagg 15900  
cgctgatag aagcggggag tttgcttac gcaagagggc gaggacgggc gagatcgcca 15960  
tcccgtgga cacttccaac cccaccccc gtctgaaacc cgtgacgctg caacaggtgt 16020  
tgccggtgag cccccctcgc cgacgcggca taaaacgcga gggcggcgag ctgcagccca 16080  
ccatgcagct cctggttccc aagaggcaga aactagagga cgtactggac atgataaaaa 16140  
tgagcccca cgtgcagccc gatattaaaa tccgtcccat caaagaagt ggcgccggaa 16200  
tgggcgtgca gaccgtggac atccagatc ccatgaccag cgcgcacag gcggtagagg 16260  
ccatgcagac cgacgtgggg atgatgacg acctgcccgc agctgctgcc gccgtggcca 16320  
gcgcccgcac gaaaaccgaa gccggcatgc agaccgacc gtggacggag gcgcccgtgc 16380  
agccggccag aagacgcgtc agacggacgt acggccccgt ttctggcata atgcccgggt 16440

---

-continued

---

acgcgctgca tccttccatc atccccaccc cggctaccg ggggcgcacc taccgtccgc 16500  
gacgcagcac cactcgcgcg cgctcgcgca cggcacgagt cggcaccgcc agagtgcgac 16560  
gcgtaacgac acgtcgcgcg cgcgcgttga ccctgcccgt ggtgcgctac catcccagca 16620  
ttctttaaaa aaccgctcct acgttgacaga tgggcaagct tacttgctga ctccgtatgg 16680  
ccgtgcccgg ctaccgagga agatcccgcc gacgacggac tttgggaggc agcggtttgc 16740  
gccgcgctcg ggcggttcac cgggcctca agggaggcat tctgccggcc ctgatcccca 16800  
taatcggcgc agccatcggg gccattcccg gaatcgccag cgtagcggtg caggctagcc 16860  
agcgcactg attttactaa ccctgtcggg cgcgcctct ctttcggcag actcaacgcc 16920  
cagcatgga gacatcaatt tctcctctct ggcccccg caccgcaacg gcccgatat 16980  
ggggacgtg agcagatcgc gcacgaacca gatgaacggg ggcgctttca attggagcgg 17040  
tgtgtggagc ggcttgaaaa atttcggttc cactctgaaa acttacggca accgggtgtg 17100  
gaactccagc acggggcaga tgctgagga caagctaaag gacacgcagt ttcagcaaaa 17160  
ggtggtggac ggcacgcctt cgggcctcaa cgggcgcgct gacctggcca accagcccat 17220  
tcaaaaggaa ataacagcc gcctggagcc cggcccgag gtggaggaga acctgcccc 17280  
tctggaggcg ctgccccca agggagagaa cgcgcccg cccgacatgg aggagacgct 17340  
agttactaag agcagaggc cgccatcata cgaggaggcg gtgggtagct cgcagctgcc 17400  
gtccctcacg ctgaagccca ccacctatcc catgaccaag cccatcgcct ccatggcgcg 17460  
ccccgtgga gtgcacccgc ccacgcagc ggtggccact ttggacctgc cgcgccccga 17520  
accggcaac cgcgtgcctc ccgtcccat cgctccgcg gtttctcgc ccgccatccg 17580  
ccccgtcgc gtggcactc cccgctatcc gagccgaac gccaaactggc agaccaccct 17640  
caacagtatt gtcggactgg gggggaagtc tctgaagcgc cgtcgtggt tttaaagcac 17700  
aatattata acgagtagcc ctgtcttaat ccacgttgt atgtgtcct atatcacgcg 17760  
ttcagagcct gaccgtcct caagatggc actccgtcga tgatgccga gtggtcgtac 17820  
atgcacatcg ccgggcagga cgcctcggag tacctgagcc cgggtctggt gcagtttgc 17880  
cgtgcgacg aaactactt ctactgggc aacaagttca ggaacccac cgtggcgc 17940  
accacgacg tcaccaccga tcggtcccag cgactgaca tccgcttctg cccgtggac 18000  
aaggagaca ccgcttact ctacaaaacc cgcttcacgc tggccgtggg cgacaaccg 18060  
gtgctagaca tggccagtac ctactttgac atccgcgcg tgatcgaccg cggacctagc 18120  
ttcaagcctt actccggcag gcttacaac tcaactggct ccaaggggc gcccaacaac 18180  
agccaatgga acgcccagga taacgggaac aagccagtgt gttttgctca ggcagctttt 18240  
ataggtcaaa gcattacaaa agacggagt caaatacaga actcagaaaa tcaacaggct 18300  
gctgcccaca aaacttacca accagagcct caaattggag tttccacctg ggataccaac 18360  
gttaccagta acgctgcccg acgagtgta aaagccacca ctcccagct gccatgttac 18420  
ggttcatacg ccaatcccac taatccaaac ggggtcagg caaaaacaga aggagacatt 18480  
tcgctaaact ttttcacaac aactgoggca gcagacaata atcccaaagt gttcttttac 18540  
agcgaagatg taaaccttca agccccgat actcacttag tatataagcc aacggtggga 18600  
gaaaacgtta tcgcccagga agccctgcta acgacgagc cgtgtcccaa cagagcaaac 18660  
tacataggtt tccgagataa ctttatcggg ttaatgtatt ataacagcac agggaacatg 18720

---

-continued

---

ggagttctgg caggtcaggc ctgcagtta aacgcagttg tagacctgca agatcgaaac 18780  
acggaaactgt cctatcagct aatgctagat gctctgggtg acagaactcg atatttctca 18840  
atgtggaatc aggccgtgga cagctacgat ccagacgtta ggattatcga gaacctggg 18900  
gtggaagacg agctgccaa ttactgtttt ccaactcccag gcatgggtat ttttaactcc 18960  
tacaaggggg taaaaccaca aaatggcggg aatggtaact gggagcaaa cggggaccta 19020  
tcaaatgcca atgagatcgc tttaggaac attttgcca tggaaattaa cctccacgca 19080  
aaactgtggc gcagcttctt gtacagcaat gtggcgctgt acctgccaga cagctataaa 19140  
ttcactcccg ctaacatcac tctgcccgcc aaccaaaca cctacgagta tatcaacggg 19200  
cgcgtcactt ctccaacct ggtggacacc tttgttaaca ttggagccg atggtcgccc 19260  
gatcccattg acaacgtcaa cccctttaac catcacggga acgcgggcct cegttaccgc 19320  
tccatgctgc tgggaaatgg acgctgtgtg cctttccaca tacaagtgc gcaaaaattt 19380  
ttcgcgatta agaacctcct gcttttgccc ggctcctaca cttacgagtg gagcttcaga 19440  
aaagacgtga acatgattct gcagagcacc ctgggcaatg atcttcgagt ggacggggcc 19500  
agcgtccgca ttgacagcgt caactgtac gccactttt tccccatggc gcacaacacc 19560  
gcttctacct tgggaagccat gctgcgaaac gacaccaacg accagtcggt taacgactac 19620  
ctcagcgccg ccaacatgct ttatcccatt ccggccaacg ccaccaacgt tcccatttcc 19680  
attccctccc gaaactgggc ggcttccgg ggatggagct tcaccgcct taaagccaag 19740  
gaaacgcctt ccttgggctc cggctttgac ccctactttg tgtactcagg caccattcct 19800  
tacctggacg gcagctttha cctcaaccac actttcaaac gtctgtccat catgttcgat 19860  
tcttccgtaa gttggccggg caacgaccgc ctctgacgc cgaacgagtt cgaaaattaag 19920  
cgcatgtgg acggggaagg ctacaacgtg gctcaaagta acatgaccaa agactggttt 19980  
ttaattcaaa tgctcagcca ctacaacatc ggctaccaag gcttctatgt tcccgagggc 20040  
tacaaggatc ggatgtatct tttctccga aactttcagc ccatgagccg ccaggtgccc 20100  
gatcccaccg ctgccggcta tcaagccgtt cccctgccc aacaacaaa caactcgggc 20160  
ttgtgggggt acatggggcc gacctgccc gaaggacagc catacccggc caactacccc 20220  
tatcccctga tcggcgctac cgccgtcccc gccattacc agaaaaagt tttgtcgcac 20280  
cgcgcatgt ggcgatacc tttttccagc aactttatgt caatgggggc cctgaccgac 20340  
ctcggacaga acatgcttha cgctaactcc gccatgccc tggatatgac ttttgagggtg 20400  
gacccatga acgagccac gttgtgtac atgctttttg aggtgttoga cgtggtcaga 20460  
gtgcaccagc cgcaccgccc tattatcgag gccgtgtacc tgcgcacccc cttctctgcy 20520  
ggcaatgcca ccacataagc cgctgaacta gctggttttt accccagatc ccatgggctc 20580  
cacggaagac gaactcgggg ccattgtgcy agacctgggc tgcggacct acttctggy 20640  
cacctttgac aagcggtttc ccgggttcgt gtctcctcgc aaactcgcgt gcgcatcgt 20700  
gaataccgcc ggccgagaga ccggaggaga gcattggcta gctctgggct ggaaccccc 20760  
ctcgtccacg tttttctgt tcgacccctt tggcttttca gaccaacgt tgaagcagat 20820  
ctatgcattt gaatatgag gtctactcaa gcgaagcgc ctggcctcct ccgccatca 20880  
ctgtctaacc ctggtaaaga gcaactcagc ggttcagggc cctcacagc ccgcctgtgg 20940  
cctttttgt tgcattgttt tgcacgcctt tgtgaactgg ccggacaccc ccatggaaaa 21000

---

-continued

---

caacccacc atggacctcc tgactggcgt tcccaactcc atgctccaaa gccccagcgt 21060  
gcagaccacc ctctccaaa accagaaaa tctgtacgcc tttctgcaca agcactctcc 21120  
ctactttcgc cgccatcggg aacaaataga aaatgcaacc gcgtttaaca aaactctgta 21180  
acgtttaata aatgaacttt ttattgaact ggaaaacggg tttgtgattt ttaaaaatca 21240  
aaggggttga gctggacatc catgtgggag gccggaaggg tgggtttctt gtactggtac 21300  
ttgggcagcc acttaaactc tggaatcaca aacttgggca gcggtatttc tgggaagtgt 21360  
tcgtgcaca gctggcgggt cagctgaagt gcctgcagaa catcgggggc ggagatcttg 21420  
aagtgcagc ttatctggtt cacggcacgc gcgttgcggt acatgggatt ggcacactga 21480  
aacaccagca ggctgggatt cttgatgcta gccagggcca cggcgtcggg cacgtcaccg 21540  
gtgtcttcta tgttgacag cgaaaaaggc gtgactttgc aaagctggcg tcccgcgca 21600  
ggcacgcaat ctcccaggta gtgactca cagcggatgg gcagaagaag atgcttgtgg 21660  
ccgcggtca ttagggata ggccgctgcc ataaaagctt cgatctgctt gaaagcctgc 21720  
ttggccttgt gccctcgggt ataaaaaca ccgaggact tgttgaaaa ggtattactg 21780  
gcgcaagcgg catcgtgaaa gcaagcgcgt gcgtcttctg ttcgtaactg caccacgctg 21840  
cggccccacc gttctgaat caccttgcc ctgcccgggt tttccttgag agcgcgctgg 21900  
ccggcttcgc tgcccacatc catttccacg acatgctcct tgttaatcat gcccagaccg 21960  
tggaggcagc gcagctcctc gtcacgtcgc gtgcagtgat gctcccacac gacgcagcca 22020  
gtgggtccc acttgggctt ggaggcctcg gcaatgccag aatacaggag aacgtagtgg 22080  
tgcaaaaaac gtcccacatc ggtgcccagg gttttctggc tgctgaagggt catcgggcag 22140  
tacctccagt cctcgttaa gcaagtgttg cagatcttcc tgaagaccgt gtactgatcg 22200  
ggcataaagt ggaactcatt gcgctcggtc ttgtcgtatc tatacttttc catcagacta 22260  
tgcataatct ccatgccctt tcccaggcg caaacaatct tgggtctaca cgggtaggt 22320  
atggcctaa ggttggcct ctgaggcggc gcttgttctt cctcttgagc cctctcccga 22380  
ctgacggggg ttgaaagagg gtgcccttg gggaacggct tgaacacggg ctggcccag 22440  
gcgtcccga gaatctgcat cgggggattg ctggccgtca tggcgtgat ctgaccccgg 22500  
ggctcctca cttcgtcctc ctgaggactt tcctcgtgct tttcggggga cggtagcggg 22560  
gtagggggaa gagcgcggcg cgccttctc ttggcggca gttccggagc ctgctcttga 22620  
cgactggcca ttgtcttctc ctaggcaaga aaaacaagat ggaagactct ttctcctcct 22680  
cctcgtcaac gtcagaaagc gactcttcca cottaagcgc cgagaactcc cagcgcatag 22740  
aatccgatgt gggctacgag actcccccg cgaacttttc gccgcccc ataacaacta 22800  
acgggtggac ggactacctg gccctaggag acgtactgct gaagcacatc aggcggcaga 22860  
gcgttatcgt gcaagatgct ctcacogagc gactcgcggt tccgctggaa gtggcggaac 22920  
ttagcggcgc ctacgagcga accctcttct ccccaaagac tcccccaag aggcaggcta 22980  
acggcacctg cgagcctaac cctcactca acttctacc tgcctttgcc gtgcccagag 23040  
tactggctac gtaccacatt tttttocaaa accacaaaat ccctctctcg tgcgcgcca 23100  
accgcaccaa agccgatcgc gtgctgcgac tggaggaagg ggctcgcata cctgagattg 23160  
cgtgtctgga ggaagtccca aaaatctttg aaggtctggg ccgacgacga aagcagcag 23220  
caaacgctct ggaagagaac gcagagatc acaacagcgc cttggtagaa ctcgaggcgg 23280

---

-continued

---

acaacgccag actggccgct ctcaaaccgt ccatagaagt cacgcacttc gcctaccccg 23340  
ccgttaacct ccctccaaaa gttatgacag cggtcacatga ctcgctgctc ataaagcgcg 23400  
ctcagccctt agaccagag cacgaaaaca acagtgcga aggaaaaccg gtggtttctg 23460  
atgaggagtt gagcaagtgg ctgtcctcca acgaccccg caggttgag gaacgaagaa 23520  
aaacatgat ggccgtggtg ctagtaccg tgcaattaga atgtctgcag aggttctttt 23580  
cccaccaga gacctgaga aaagtggagg aaacgctgca ctacacattt aggcacggct 23640  
acgtgaagca agcctgcaag atttccaacg tagaacttag caacctcctc tcctacctgg 23700  
ggatcttgca cgaaaaccgc ctccgacaaa acgtgctgca cagcacactg aaaggagaag 23760  
cccgcgaga ctatgtgcga gactgcgtgt tcctagcgtc agtgtacacc tggcagagcg 23820  
gaatgggagt ctggcagcag tgcctggagg acgaaaacct caaagagctt gaaaagctgc 23880  
tggtgcgctc cagaaggcca ctgtggacca gttttgacga gcgcaccgcc gcgcgagacc 23940  
tagctgatat tatttttctt cccaagctgg tgcagactct ccgggaagga ctgccagatt 24000  
ttatgagtca aagcatcttg caaaacttcc gctctttcat cttggaacgc tcgggaatct 24060  
tgcccgccac tagctgcgcc ctaccacag atttgtgccc tctccactac cgcgaatgcc 24120  
caccgcgctg gtggccgtac acttacttgc ttaactggc caactttcta atgttccact 24180  
ctgacctggc agaagacgtt agcggcaggg ggctgctaga atgccactgc cgctgcaacc 24240  
tgtgcacccc ccaccgctct ctagtatgca aactcccct gctcaatgag acccagatca 24300  
tcggtacctt tgaatccag ggaccctccg acgcggaaaa cggcaagcag gggctctggc 24360  
taaaactcac agccgactg tggacctccg cctacttgcg caaatttcta ccagaagact 24420  
atcacgccc ccaaatataa ttttacgaaa accaatcaaa accacccaaa agcgagttaa 24480  
cggcttgctg cattacgcag agcagcatag ttgggcagtt gcaagccatt aacaaagcgc 24540  
ggcaagagtt tctcctaaaa aaaggaaaag gggcttactt ggacccccag accggcaggg 24600  
aactcaacgg acctcctca gtgcaggtt gtgtgcccc tgcgccccaa aaagaacacc 24660  
tcgcagtgga acatgccaga gacggaggaa gaggagtgga gcagtgtgag caacagcgaa 24720  
acggaggaag agccgtggcc cgagggtgac aacggggaag aggacacgga gggacggcga 24780  
agtcttcgcc gaagaactct cgcgctgcc cccgaagtc cagccggccg cctcgccca 24840  
agatcccga cacaccgta gatgggatag caagacaaa aagccgggta agagaacgc 24900  
tcgccccgc cagggctacc gctcgtggag aaagcacaaa aactgcatct tatcgtgctt 24960  
gctccagtgc ggcggagacg tttcgttcac cgttagatag ttgcttttta acaaggggt 25020  
ggcgtcccc cgtaactgcc tccactacta cgtcactct tacagctccg aagcggacgg 25080  
ctaagaaaac gcagcagttg ccggcgggag gactgcgtct cagcgcccga gaacccccag 25140  
ccaccagga gctccgaaac cgcataattc ccacctcta cgtatcttt cagcaaagcc 25200  
gggggcagca gcaagaactg aaaataaaaa accgcacgct gaggtcgtt acccgaagct 25260  
gcctctatca caagagcgaa gacgagctgc agcgaacct ggaggacgca gaagcgtgt 25320  
tccagaagta ctgcccagacc accctaaata actaaaaag cccgcgcgcg gaacttcaa 25380  
ccgtctgacg tcaccagccg cgcgcaaaa tgagcaaaga gattcccacg ccttacatgt 25440  
ggagttacca gccgcagatg ggattagccg ccggcgcgcg ccaggattac tccacgaaaa 25500  
tgaactggct cagcgcgggg ccccaocatga tttccgcgt aaacgacatt cgcgccacc 25560

---

-continued

---

gcaatcagct attgtagaa caggetgctc tgaccgccac gccccgtaat aacctgaacc 25620  
ctcccagctg gccagctgcc ctggtgtacc aggaaacgcc tccaccacc agcgtacttt 25680  
tgccccgtga cccccaggcg gaagtccaga tgactaacgc gggcgcgcaa ttagcggcg 25740  
gatccccggtt tcggtacaga gttcacggcg ccgcacccta tagcccaggt ataaagaggc 25800  
tgatcattcg aggcagaggt gtccagctca acgacgagac agtgagctct tcgcttggtc 25860  
tacgaccaga cggagtgttc cagctcgcgg gctcgggccc ctcttcgttc acgcctcgcc 25920  
aggcatacct gactctgcag agctctgcct ctcagcctcg ctcgggagga atcggacccc 25980  
ttcagtttgt ggaggagttt gtgccctcgg tctactttca gcctttctcc ggatcgcccg 26040  
gccagtaccg ggacgagttc atccccaaact tcgacgcggg gagtgactct gtggacggtt 26100  
atgactgatg tcgagcccgc ttcagtgcga gtggaacaag cgcggctcaa tcacctgggt 26160  
cgttgccgcc gccgctgctg cgtggctcgc gacttgagct tagctctcaa gtttgtaaaa 26220  
aaccgcctcg aaaccgggag cgctgtgcac gggttggagc tagtgggtcc tgagaaggcc 26280  
accatccacg ttctcagaaa ctttggtaa aaaccattt tggtaaacg agatcagggg 26340  
ccttttgtaa tcagcttact ctgcacctgt aaccatgttg accttcacga ctattttatg 26400  
gatcatttgt gcgctgaatt caataagtaa agcgaattct taccaagatt atgatgtcca 26460  
tgactgttcc tcgccactat acgatgttgt gccagtaaac tctctgtcgc acatctatct 26520  
gaaactgttcc ttttggctcg cacagcttac ttggtactac ggtgacaccg tcctttctgg 26580  
ctcactgggc agctcacacg gaataaacct tcacctcttt tcgccgtttc gatacggaaa 26640  
ctacagctgt cgtgccggta cctgcctcca cgttttcaat cttcagccct gtccaccgac 26700  
caaacttgta tttgtcgact ctaagcactt acagctcaac tgcagcattc taggccccag 26760  
tatcttgggg acatacaata aaatcaggtt ggtggaattt gtctactacc caccacgcg 26820  
ccgcggtttt ggggaaattc ctttccagat ctactacaac tatcttgcca cacattatgc 26880  
aagtcaacag caactaaact tgcaagcacc cttcacgcca ggagagtact cctgtcacgt 26940  
aggctcctgc acagaaactt ttattctctt caacagatct tctgccattg aacgcttcac 27000  
tactaactac ttttagaaacc aagttgtgct tttcactgac gaaacccta acgtcacctc 27060  
ggactgtgca tgtttttctc atgacaccgt aacttgact cttacaata ctctctggct 27120  
cgcgttcgat aaccaagct tgattgtaa aaattttgat ttaacctta ctaaaccctc 27180  
tcctcgcaa atagttatct ttgctcctt taatccaaa actacctag cctgtcaggt 27240  
ttgtttaaag ccttgccaaa caaactttaa gtttgtttat ttgctcgcg aatctgcaa 27300  
actcatagaa aaatacaaca aagcggcctt cttggctcct aaaacctct accactggct 27360  
aacctacacg gggctgtttg cactaattgt tttttccta attaacattt ttatatgttt 27420  
cttgcttcc tccttctttt cgcgaacacc gttgccgag aaagacctct ccttattact 27480  
gtagcgttg ctatacaaaa ccaagagtgg tcaaccgtgc tctcaatcta tttcaattt 27540  
ttcattttgt ccttaatact ttctcttatt gtcgttaaca atgatctgga gcattgtct 27600  
cgcctttttt tggtgtcta gtcaaaaagc cactattttt cacaggtatg tggagaagg 27660  
aactagcacc ctctttacga tacctgaaac aattaaggcg gctgatgaag tttcttggt 27720  
caaaggctcg ctctcagacg gcaaccactc attctcagga cagacccttt gcatccaaga 27780  
aacttatttt aaatcagaac tacaatacag ctgcataaaa aactttttcc atctctacaa 27840

---

-continued

---

catctcaaaa ccctatgagg gtatttacia tgccaagggt tcagacaact ccagcacacg 27900  
gaacttttac tttaatctga cagttattaa agcaatttcc attcctatct gtgagtttag 27960  
ctcccagttt ctttctgaaa cctactgttt aattactata aactgcaacta aaaatcgctt 28020  
tcacaccacc ataatctaca atcacacaca atcaccttgg gttttaaacc taaaattttc 28080  
tccacacatg ctttcgcaat ttctcacgca agttaccgtc tctaacataa gcaagcagtt 28140  
tggcttttac tatcctttcc acgaactgtg cgaaataatt gaagccgaat atgaaccaga 28200  
ctactttact tacattgcca ttggtgtaat cgttgtttgc ctttgctttg ttattggggg 28260  
gtgtgtttat ttgtacattc agagaaaaat attgctctcg ctgtgctcct gcggttacia 28320  
agcagaagaa agaattaaaa tctctacact ttattaatgt tttccagaaa tggcaaaact 28380  
aacgctccta cttttgcttc tcacgccggt gacgcttttt accatcactt tttctgccgc 28440  
cgccacactc gaacctcaat gtttgccacc ggttgaagtc tactttgtct acgtgttgct 28500  
gtgctgcggt agcgtttgca gtataacatg ttttaccttt gttttcttc agtgcattga 28560  
ctacttctgg gtcagactct actaccgag acacgcgctt cagtatcaaa atcaacaaat 28620  
tgccagacta ctcggtctgc catgattgtc ttgtatttta ccctgatttt ttttcactt 28680  
acttgcgctt gtgattttca cttcactcaa ttttgaaaa cgcaatgctt cgacccgcgc 28740  
ctctccaacg actggatgat ggctcttgca attgccacgc ttggggcggt tggacttttt 28800  
agtgtttttg ctttgcatca caaatttaag actccatgga cacatggctt tctttcagat 28860  
tttccagtta cacctactcc gccgcctccc cgggccatcg acgtgcctca ggttccctca 28920  
ccttctccat ctgtctgcag ctactttcat ctgtaatggc cgacctagaa tttgacggag 28980  
tgcaatctga gcaaagggtc atacacttcc aacgccagtc ggaccgcgaa cgcaaaaaca 29040  
gagagctgca aaccatacaa aacacccacc aatgtaaacy cgggatattt tgtattgtaa 29100  
aacaagctaa gctccactac gagcttctat ctggcaacga ccacgagctc caatacgtgg 29160  
tcgatcagca gcgtcaaac tgtgtattct taattggagt tccccctatt aaagttactc 29220  
aaaccaaggg tgaaccaag ggaaccataa ggtgctcatg tcacctgtca gaatgccttt 29280  
acaacttagt taaaacccta tgtggcttac atgattctat cccctttaat taaataaact 29340  
tactttaaat ctgcaatcac ttctctgtcc ttgtttttgt cgccatccag cagcaccacc 29400  
ttccctctt cccaacttcc atagcatatt ttccgaaaag aggcgtactt tcgccacacc 29460  
ttaaagggaa cgtttacttc gctttcaagc tctcccacga ttttcattgc agatagaaa 29520  
cgcgcctaaag tggagaag atttaacccc gtttatccct atggatattc tactccgact 29580  
gacgtggctc ctccctttgt agcctctgac ggtcttcaag aaaaccacc tggggtcttg 29640  
tccttaaaaa tatccaaacc tttactttt aatgcctcca aggccttaag cctggctatt 29700  
ggtccaggat taaaattca agatggtaaa ctagtggggg agggacaagc aattcttgca 29760  
aacctgccgc ttcaaatcac caacaacaca attcactac gttttgggaa cacactgcc 29820  
ttgatgaca ataatgaact ccaaaccaca ctaaaatctt catcgccct taaaatcaca 29880  
gaccagactc tgtcccttaa cataggggac agccttgcaa ttaagatga caaactagaa 29940  
agcgtcttc aagcgacct cccactctcc attagcaaca acaccatcag cctcaacgtg 30000  
ggcaccggac tcaccataaa tggaaacgtt ttacaagctg ttcccttaaa tgetctaagt 30060  
cccctaacta tttcaacaa taacatcagc ctgctctatg gcagttcct gacggtgctt 30120

---

-continued

---

aacaatgaac tgcaaagcaa cctcacagtt cactccccctt taaaactcaa ctccaacaac 30180  
tcaatttctc tcaacactct atctccggtt agaatcgaga atggtttcct cacgctctat 30240  
ttgggaacaa aatctggctt gctagttcaa aacagtggct taaaagtca agcgggctac 30300  
ggcctgcaag taacagacac caatgctctc acattaagat atctcgctcc actgaccatt 30360  
ccagactcgg gctcagaaca aggcattctt aaagtaaaca ctggacaggg cctaagtgtg 30420  
aaccaagctg gagcgcttga aacatcccta ggaggtggat taaaatagc tgataacaaa 30480  
ataacctttg atacaggaaa cggactgaca ttatctgaaa ataaacttgc agtagctgca 30540  
ggtagtggtc taacttttag agatgggtgcc ttggtagcca cgggaaccgc atttacgcaa 30600  
acactgtgga ctacggctga tccgtctccc aactgcacaa ttatacagga ccgacacaca 30660  
aaatttactt tggcgcttac cattagtggg agccaagtgc tggggacggg ttccattatt 30720  
ggagtaaaa gcccccttc aagtagcata ccgtcagcta ccgttacagt acaacttaac 30780  
tttgattcca acggagccct attgagctcc tcttactta aaggttactg ggggtatcgc 30840  
caaggtccct caattgacc ttacccata attaagcct taaactttat gccaaactca 30900  
ctggcttate ccccgggaca agaaatcaa gcaaatgta acatgtacgt ttctactttt 30960  
ttacagggaa atccacaaag accaatagtt ttaaacatca cttttaataa tcaaaccagc 31020  
gggttttcca ttagatttac atggacaaat ttaaccacag gagaagcatt tgcaatgccc 31080  
ccatgcactt tttcctacat tgtgacaaa caataaacta tgtaaccctc accgttaacc 31140  
cgctccgcc cttccatttt atttataaaa ccaccgac caccctttca gcagtaaaaa 31200  
attgcatgtc agtaggggca gtaaaacttt tgggagttaa aatccacaca ggttcttcac 31260  
aagctaagcg aaaatcagtt acacttataa aaccatcgct aacatcggac aaagacaagc 31320  
atgagtccaa agcttccggt tctggatcag atttttgttc attaacagcg ggagaaacag 31380  
cttctggagg attttccatc tccatctcct tcacagttc caccatgtcc accgtggtca 31440  
tctgggacga gaacgacagt tgtcatacac ctcataagtc accggtcgat gacgaacgta 31500  
cagatctcga agaatgtcct gtcgcccct ttcggcagca ctgggccgaa ggcgaaagcg 31560  
cccatgttta acaatggcca gcaccgccg cttcatcagg cgcctagttc ttttagcgca 31620  
acagcgcagc cgcagctcgc taagactggc gcaagaaaca cagcacagaa ccaccagatt 31680  
gttcatgac ccataagcgt gctgacacca gccatacta acaaatgtt tcaactattct 31740  
agcatgaatg tcatatctga tgttcaagta aattaaatgg cgcacctta tgtaaacact 31800  
tcccacgtac aacacctcct ttggcatctg ataattaacc acctcccgat accaaatata 31860  
tctctgatta atagtcgccc cgtacactac ccgattaaac caagttgcca acataatccc 31920  
ccctgcata cactgcaaag aacctggagc gctacaatga cagtgcaaag tccacacctc 31980  
gttgccatgg ataactgagg aacgcottaa gtcaatagtg gcacaactaa tacaacatg 32040  
taaatagtg ttcaacaagt gccactcgtg tgaggtgagt atcatgtccc agggaacggg 32100  
ccactccata aacactgcaa aaccaacaca tctaccatc ccccgacgg cactcacatc 32160  
gtgcatgggt tccatatac agtccggaag ctgaggacaa ggaaaagtct cgggagcatt 32220  
ttcatagggc gtagtgggt actccttga ggggttcagt cggcaccggg atctcctcac 32280  
cttctgggcc ataacacaca agttgagatc tgatttcaag gtactttctg aatgaaaacc 32340  
aagtgccttc ccaacaatgt atccgatgct ttcggtcccc gcgtcggtag cgtccttgc 32400



-continued

---

```

agtacacacg gaacaaccac tcacgcagcg ccagaagaca gttttccgcg gacggtgaca 32460
agttaatccc cctcagtctc agagccaata tagtttcttc cacagtagca taggccaac 32520
ccaaccagga aacacaagct ggcacgtccc gttcaacggg aggacaagga agcagaggca 32580
gaggcatagg caaagcaaca gaatttttat tccaactggt cacgtagcac ttcaaacacc 32640
aggtcacgta aatggcagcg atcttggggt tcctgatgga acataacagc aagatcaaac 32700
atgagacgat tctcaagggt attaaccaca gctggaatta aatcctccac gcgcacattt 32760
agaaacacca gcaatacaaa agcccgggtt tctccgggat ctatcatagc agcacagtca 32820
tcaattagtc ccaagtaatt ttcccgttcc caatctgtta taatttgagc aataatgccc 32880
tgtaaatcca agccggccat ggcgaaaagc tcagataatg cactttccac gtgcattcgt 32940
aaacacaccc tcatctgtgc aatccaaaaa gtcttcttct tgagaaacct gtagtaaat 33000
aagaatcgcc aggttaggct cgatgcctac atcccggagc ttcattctca gcatgcactg 33060
caaatgatcc agcagatcag aacagcaatt agcagccagc tcatcccggg tttccagttc 33120
cggagttccc acggcaatta tcaactgaaa cgtgggacaa atcgaaataa catgagctcc 33180
cacgtgagca aaagccgtag gccacgtgca ataatcacag aaccagcggg aaaaagattg 33240
cagctcatgt ttcaaaaagc tctgcagatc aaaattcagc tcatgcaaat aacacagtaa 33300
agtttgcggt atagtaaccg aaaaccacac gggtcgacgt tcaaacatct cggcttacct 33360
aaaaaagaag cacattttta aaccacagtc gtttctgaa caggaggaaa tatggtgagg 33420
cgtaaaacca gacgcgccac cggatctccg gcagagccct gataatacag ccagctgtgg 33480
ttaaacagca aaacctttaa ttcggcaacg gttgaggtct ccacataatc agcggcccaca 33540
aaaatcccat ctgcaacttg ctgcgtagg gagctaaaat gccagctata gcccctggc 33600
acccgaacgc taatctgcaa gtatatgaga gccaccccat tcggcgggat cacaaaatca 33660
gtcggagaaa acaacgtata cccccggac tgcaaaaagct gttcaggcaa acgcccctgc 33720
ggtcctctc ggtacaccag caaagcctcg ggtaaagcag ccatgccaag cgcttaccgt 33780
gccaagagcg actcagacga aaaagtgtac tgaggcgctc agagcagcgg ctatatactc 33840
tacctgtgac gtcaagaacc gaaagtcaaa agttcaccg gcgcgccga aaaaaccgc 33900
gaaaatccac ccaaaaagcc cgcgaaaaac acttccgtat aaaatttccg ggttaccggc 33960
gcgtcaccgc cgcgcgacac gcccgccccg ccccgcgctc ctccccgaaa cccgccgcgc 34020
ccacttccgc gttccaaga caaaggtcgc gtaactcgc ccacctcatt tgcattgtaa 34080
ctcggtcgcc atcttgcggt gttatattga tgatg 34115

```

```

<210> SEQ ID NO 7
<211> LENGTH: 44
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: primer P5L

```

```

<400> SEQUENCE: 7

```

```

gcgcacgcgt ctctatcgat gaattccatt ggtgatggac atgc 44

```

```

<210> SEQ ID NO 8
<211> LENGTH: 36
<212> TYPE: DNA
<213> ORGANISM: Artificial

```

-continued

---

```

<220> FEATURE:
<223> OTHER INFORMATION: primer P5ITR

<400> SEQUENCE: 8

gcgcatthaa atcatcatca ataataacc tcaaac          36

<210> SEQ ID NO 9
<211> LENGTH: 31
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: primer P5XTOP

<400> SEQUENCE: 9

gatacctagg aacgaggagg atttgatatt g          31

<210> SEQ ID NO 10
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: primer P5XBOT

<400> SEQUENCE: 10

atgtacgcct cgcgctcac          20

<210> SEQ ID NO 11
<211> LENGTH: 31
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: primer P5E4

<400> SEQUENCE: 11

gatcgaattc ccactctgta ccccatctct g          31

<210> SEQ ID NO 12
<211> LENGTH: 31967
<212> TYPE: DNA
<213> ORGANISM: Simian adenovirus
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (13796)..(15322)
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (18257)..(21010)
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (27192)..(29015)

<400> SEQUENCE: 12

catcatcata atatacctta ttgggaacg gtgccaatat gataatgagg aggcgggggtt          60
aggggtggag tgaggggtgg gtgaggatga cgcgggcgcg gggcggggtg ggagtctgac          120
gtggggcgcg ggggtgagcg cgagggtgag ggcggggcga gggcggcggg cgcggcggaa          180
ttgagctaca cggtagtaag tttagcgga aattaagtga attgggcgtg ttttttghaa          240
ctttttgacg tacacggtag taagtttgag cggaaattaa gtgaattggg cgtgtttttt          300
gtaacttttt gacgtacacg gtagtaagtt tgagcggaaa ttaagtgaat tgggcgtggt          360
ttttgtaact ttttgacgta cacggtagta agtttgagcg gaaattaagt gaattgggcg          420
tgttttttgt aactttttga cgtacacggt agtaagtttg agcggaaatt aagtgaattg          480

```

-continued

---

|   |      |
|---|------|
| ggcgtgtttt ttgtaacttt ttgacgtaca cggtagtaag tttgagcggg aattaagtga | 540  |
| attgggcgtg ttttttgtaa ctttttgacg tacacggtag taagtttgag cggaaattaa | 600  |
| gtgaattggg cgtgtttttt gtaacttttt ggtcattttg gcgcgaaaac tgagtaatga | 660  |
| ggaagtgaga cggactctgc ccttttttac ggttgggagg gaaaactgct gatcagcgct | 720  |
| gaactttggg ctctgacgcg gtggtttccc tacgtggcag tgccacgaga aggctcaaag | 780  |
| tcctcgtttt attgtgtgct cagccttttt gagggatatt aaacaccgtc agaccgtcaa | 840  |
| gaggccactc ttgagtgaga gcgagtagag ttttctctc cgtcgctgcc gcggctgctc  | 900  |
| agtcttaccg ccaggatgcg aatgctgccg gagatcttca ccgggtcctg ggaagatggt | 960  |
| ttccagggac ttttagaatc tgaagacaac tttccccaac ctctgagcc ggaggagcta  | 1020 |
| cctgagggtt cgcttcacga tctgtttgac gtggagggtg agagccccga cggagatccg | 1080 |
| aacgaggaag ctggtgatgg tatgttcccc gactggatga tatctcagag cgagagtgct | 1140 |
| gaaggcagtg cggactcggg cgtttctggy gttggaaacc tggtgagggt ggatctggac | 1200 |
| ttgaagtgtt acgaggaagg ttttctctc agcgactcag agactgatga agcctcagaa  | 1260 |
| gcggaaggtc aagaggagtc tgtgtgtggt tatgtgaaga ttaatgaggg ggagaacctg | 1320 |
| ctggtgttg actgtccgga ccaacctgga catggctgtc gagcctgtga ctttaccgg   | 1380 |
| gggaccagcg gaaaccggga agctatctgt gctttgtgct acatgcgtct gaacgagcac | 1440 |
| tgcatataca gtgagtgta ttcattgggtt atttatggg aaagttggg gaaagtcttg   | 1500 |
| agaaggggaa aagtttaaca tgtcattttt gtacttgata ggtccagttt cggacgctga | 1560 |
| gggggattct gaggccctg ctggtccttc ccagccctca ccctgctctt tgaccgccac  | 1620 |
| gcccgcacct gacctagta gaccaacgcc ctgccagtg tcctgtagac gacgtgcagc   | 1680 |
| tgtaattgc atagaagatt tattggcccc tgatgacgag aacgcacctt tgaacctgtg  | 1740 |
| cctgaaacgc cctaagacat cttgagtgtt tatgctgtta ataaaagtgt tgacccttag | 1800 |
| atcctgtgtt tattccttgg gcgtgtgccc gggatatata agcagctgcg gctggagtg  | 1860 |
| ttagtttatt ctgatggagt actggagtga gctgcagaat taccagagcc tccggcgcct | 1920 |
| gctggagtgt gcctctgcc gaacatccac ctgctggagg ttctgttttg gctogactct  | 1980 |
| cagtaacgtg gtgtatcggg tgaagcaaga gtacagctcg cgctttctg agctgttggc  | 2040 |
| ccgctaccg gctgtttttg tttctctgga tctaggccat cacgtttatt tccaagaagc  | 2100 |
| tgtagtcaga tatttgatt tttctactcc cgggcgtgcg gttctgoga ttgcttcat    | 2160 |
| ctgctttgtg ctagatgat ggagcgcaca aaccgcctg agcccgggtt acaccctgga   | 2220 |
| ctacctgacc atgtccctgt ggaggccat gctgcggaag aggaggtct caggcttctc   | 2280 |
| gcccgcgcg cctccgacg gactggatcc ggtgctggag gactcggagc tggaggagga   | 2340 |
| ggagaacccg agggccggcc tggaccctcc ggcggaatag tgacggaacc ggaggatccc | 2400 |
| caagagggta ctagttaggg gggagggggg ccgaagagaa agcgggatga agaggaggcg | 2460 |
| atggacccc acaggtttct aaaagaactg actttaagct taatgtctaa gagaagacc   | 2520 |
| gagacggtgt ggtggtctga tttggagaag gagttccacc agggggagat gaatctgtt  | 2580 |
| tacaagtatg ggtttgagca ggtgaagact cactggctgg aagcctggga ggactggag  | 2640 |
| atggctttta acatgtttgc caagtggtgc ctgcccccg aactattta cacgtgact    | 2700 |
| aagacggtgg aaatccgcaa gcctgtgtat gtgattggca acggggccgt ggttcggtt  | 2760 |

-continued

---

|   |      |
|---|------|
| cagaccaccg accgggtggc ctttaattgc tgtatgcaga acctgggccc gggggtgatt | 2820 |
| aatcttaatg gagtgacctt ttgcaatgtc agattcgcgg gggatggatt caacgggacg | 2880 |
| gtgtttgccc ccaccacca gataacccta cacggggtgt tcttcagca tgtaggcggg   | 2940 |
| gcttgtgtag atacctgggc gagggcctct gtgaggggct gcacctttgt gggctgttg  | 3000 |
| aaagcggtyg tgggtcgacc caagagtgtg ctgtctgtga agaatgtgt gttgagaga   | 3060 |
| tgtctgatgg ccatggtggt ggagggccag ggtaggatcc gccataacgc gggctccgag | 3120 |
| aatacctgtt ttgcctgct gaagggtacg gcgacctga agcataacat gatctgcggg   | 3180 |
| gtgggtcact cgcagctgct gacctgtgct gatggcaact gccaggccct gcgcacggtg | 3240 |
| catgtggtgt cccaccggcg ccgccctgg ccggtgtttg aacataacat gctgatgcgc  | 3300 |
| tgtaccatgc acctgggcta ccgccggcgt gtgtttgtgc cccatcagtg taacctgacc | 3360 |
| cacaccaag tttgtctgga gacggatgct ttttcgcgag tgaatctgaa tggggtgttc  | 3420 |
| gatctgacta tggagatgta caagatagtg agatttgatg aatcaaagac ccgttgtcgc | 3480 |
| ccctgcgagt gcggtgcca tcacctgag atgtatccc tgacctgaa cgtgacggag     | 3540 |
| gagctgcgcc cggaccacca gatgctgtcc tgtctgcgca ccgattacga aagcagcgat | 3600 |
| gaggattaag aggtgagggg cggggccttc atggggata aagggtgggg aggaggtggg  | 3660 |
| gaggggaaa acccaaatg agcggatcga tggaaaggag cgctgtgagt tttgagggcg   | 3720 |
| gggtgttcag cccatatctg acaaccctgc tccccgcctg ggcaggagtg cgtcagaatg | 3780 |
| tgggtggctc caactgggac ggacgtccgg tggccccctg caactccgcc actctcact  | 3840 |
| acgccaccgt cggatcgtcg ctggacaccg ccgctgccgc cgcgcttca gccgccgctt  | 3900 |
| ctactgctcg cggatggca gctgatttc gactgtatca gcaactggct gcgcctcgt    | 3960 |
| cgtcgtgag agaagatgat gccctgtccg tgggtctgac ccgctggag gagctgtccc   | 4020 |
| agcagctgca agagctgtct gccaaagtgg atgcacagaa cgtccccgct acccaatgaa | 4080 |
| taataaacg agacaccgag tgtgtttgga aatcaaatg tgtttttatt tgtttttct    | 4140 |
| ggcgcgtag gcccttgacc acctgtcgcg gtcgttaagg acctgtgga tgttttccag   | 4200 |
| caccocgtag aggtgggctt ggatgttag gtacatgggc atgagccgt ctcgggggtg   | 4260 |
| gaggtagcac cactggaggg cgtcgtgctc gggggtggtg ttgtagataa tccagtcgta | 4320 |
| gcagggtttt tgggcatgga agcggagat gtctttgaga agcaggctga tggccagggg  | 4380 |
| gagggccttg gtgtaggtgt tcacaaagcg gttgagctgg gagggatgca tgcgggggga | 4440 |
| gatgagatgc atcttggcct gaatcttag gttggcgtg ttgccgccca gatcccgcg    | 4500 |
| ggggctcatg ttgtgcagga ccaccaggac ggtgtagccg gtgcacttg ggaatttgc   | 4560 |
| atgcaacttg gaaggaagc cgtggaagaa cttggagacc cccttgtggc gcgagaggtt  | 4620 |
| ctccatgcat tcgtccatga tgatggcgat gggaccctg gcggccgcc tggcgaagac   | 4680 |
| gttgtcgggg tgggagacgt cgtagttctg ttccagggtg agctcgtcgt aggccatttt | 4740 |
| gacgaagcgg gggagcaggg tgcccactg ggggacgat gtacctcgg gaccocgggc    | 4800 |
| gtagttgccc tcgcagattt gcatctccca ggccttgatc tccgaggggg ggtatcatgc | 4860 |
| caactggggc gcgatgaaga agacggtctc cggggcgggg ttgatgagct gggaggagag | 4920 |
| gaggttcgcg agcagctgct acttgccgca cccggtgggc ccgtagatga cccogatgac | 4980 |
| gggttcgagc tggtagttta aggagctgca gctgccgtcc tcgcgcagga acggggcgac | 5040 |

-continued

---

|  |      |
|--|------|
| ctcgttcate atgcttctga cgtgatggtt ttccctgacg aggtcttgca agagccgctc    | 5100 |
| gccgcccagc gagagaagct cttccaggct gcggaatgc ttgaggggtt tgaggccgctc    | 5160 |
| ggccatggtc atcttttcca gggactggcg gagcaggtaac aggcgggtccc agagctcgggt | 5220 |
| gacgtgttct acggcatctc gatccagcag acttcttgggt tgcgggggtt ggggcggctt   | 5280 |
| tggtgttagg ggaccagccg gtgcgcgtcc agggaggcga gggtagcgtc tttccagggc    | 5340 |
| cgacgcgttc gcgtgagggt ggtctcggtg acggtgaagg gatgcgtcc cggttggcg      | 5400 |
| ctggccaggg tcctcttgag actcatcctg ctgggtgga agcgggcgtc ttctccctgg     | 5460 |
| gagtcggcca ggtagcattt gagcatgagg tcgtagctga gggcctcggc cgcgtggccc    | 5520 |
| ttggcgcga gcttgctttt ggagacgtgt ccgcaggcgg gacagtgcag gcacttgagg     | 5580 |
| gcgtagagct tggggccag gaagacggac tcgggggagt aggcgtcggc gccgcactga     | 5640 |
| gcgcacgtgg tctcgcactc gacgagccag gtgagctccg ggtgttggg atcaaaaacc     | 5700 |
| agctggcccc cgtgtttttt gatgcgttc ttacctcggg tctccatgag gcggcgtccg     | 5760 |
| gcttcggtga cgaagaggct gtcggtgtcg ccgtagacgg atttgagcgc gcgctgctcc    | 5820 |
| aggggaatcc cgcgatcctc cgcgtgcagg aactcggacc actctgagac gaaggcccgg    | 5880 |
| gtccacgcga ggacaaaaga ggcgatctgg gacgggtagc ggtcgttctc caccagggga    | 5940 |
| ttccacttct ccagggtgtg caggcagagg tcgtcctcct ccgcgtccat gaaggtgatt    | 6000 |
| ggcttgaag tgtatgtcac gtgaccgtcg gggtcgcgcg tgggcttata aaagggggcg     | 6060 |
| tgcccgccct ccccgctact ttcttccgca tcgctgtgga cgagatccag ctgctcgggt    | 6120 |
| gagtagggcg gctggaagcg gggcatgacc tcggcgtga ggggtcagtt ttccacgaac     | 6180 |
| gaggtggatt tgatattgac ctgtccggcg gcgatgcttt tgacggtggc ggggtccatc    | 6240 |
| tggtcagaaa agacgatctt tttgtgtcc agcttgggtg cgaacgacct gtagagggcg     | 6300 |
| ttggagagca gcttggcagat ggagcgcagg gtctggttct tctcgcggtc ggcgcgtcc    | 6360 |
| ttggcggcga tgttagctg gacgtactcg cgggccacgc agcgcattc ggggaagacg      | 6420 |
| gtggcgcgct cgtccggcag gaggcgcacg cgcagccgc ggttgtgcag ggtgatgagg     | 6480 |
| ttccacgttg tggccacctc gccgcgagtt ggctcgttgg tccagcagag gcgcccgcc     | 6540 |
| ttgcgcgagc agaagggggg caggacgtcg agctggctct ccgcgggggg gtcggcgtcg    | 6600 |
| atggtgaaga tgcccggtag cagggtggcg tcgaagtagt cgatggcgac cgcggggctc    | 6660 |
| gcgagggcgc gttcccagtc cctgaaccgc agggcgcgct cgtaggggtt gagggcgcc     | 6720 |
| ccccagggca tgggatgggt gagggccag cgtacatgc cgcagatgct gtagacgtag      | 6780 |
| aggggctcgc ggagcacgcc gaggtaggtg ggatagcagc gtcgcggcg gatgctggcg     | 6840 |
| cgacagtagt cgtacatctc gtgcgagggg gcgagaggc cgcctccgag gtcgccgcg      | 6900 |
| tgcggtctga cggcccggta ggtgacctg cggaaagtgg cgtgcgagtt ggaggagatg     | 6960 |
| gtggccgct ggaagatggt gaagctggcc tcggggagtc cgcagggcgtc gtagacgaac    | 7020 |
| tggcgtagg agtcgcgac cttctgcacg agcgcggcgg tgacgagcac gtccagggcg      | 7080 |
| cagtagtca gggctctcgc gacgaggtcg taacggggct cttgcttctt ttcccagagt     | 7140 |
| tcgcggttga ggaggtactc ctcgcgatcc ttccagtact cttcggcccg aaagccgct     | 7200 |
| tcgtccgcca gtaagaacc cagcatgtag aagcggttga cggctcggta gggacagcag     | 7260 |
| cccttctcga cgggcaggga gtaggcctgc gcggccttc tgagcagagt gtgggtgagg     | 7320 |

-continued

---

|            |            |            |            |            |             |      |
|------------|------------|------------|------------|------------|-------------|------|
| gcgaagggtg | cgcgaccat  | gaccttgagg | aactggaacc | tgaagtcggt | gtcgtcgag   | 7380 |
| gcgccccgct | cccagagccc | gtagtcggtg | cgtttctggc | tgcgggggtt | ggcagggcg   | 7440 |
| aaggtagcgt | cgttgaagag | gatcttgccg | gcgcgcggca | tgaagttgcg | ggtgatcctg  | 7500 |
| aaggccccg  | gcacgtccga | gcggttgta  | atgacctggg | ccgcgaggac | gatctcgctg  | 7560 |
| aagccgttga | tgttggtggc | gacgatgtag | agctcgacga | agcgcgggcg | cccctgcagc  | 7620 |
| ttgggggcct | tcttgagctc | ctcgtaggtg | aggcagtcgg | gcgagtagag | gcccagctcc  | 7680 |
| tgtcggggcc | attcggccac | ctgggggttg | gcttgcaaga | agccccgcca | gagctgcagg  | 7740 |
| gcgagctggg | tctggaggcg | gtcgcggtag | tcgcggaact | ttttgcccac | cgccatcttc  | 7800 |
| tcgggggtga | ccacgtagaa | ggtgcggccg | tcctggcccc | aggcgtccca | gttctgctcg  | 7860 |
| cgggcgagac | ggcagccctc | ctcgaagagg | gcctcctccc | cgagagatg  | catgactagc  | 7920 |
| atgaagggga | cgagttgctt | gccgaaggca | cccatccacg | tgtaggtctc | tacgtcgtag  | 7980 |
| gtgacgaaga | gacgttcggt | gcgaggatgc | gagccgagag | gaaagaagtt | gatctcctgc  | 8040 |
| caccagccgg | aggagtggcg | gttgacgttg | tggaagtaga | agtcacgccg | gcggaccgtg  | 8100 |
| cattcgtgct | gatatttgta | aaagcgggcg | cagtactcgc | agcgtgcac  | gctctgcaact | 8160 |
| tcctgaacga | gatgcaccgg | gcgcccgccg | accaggaggc | ggagggggca | gtccagtggg  | 8220 |
| gcttcggcgc | gctgtccttc | agcctcgtca | tgtcttctcg | cacctgcacg | ctcctgctgt  | 8280 |
| gggtggagga | cgaggggagt | gacgacgccg | cgcgagccgc | aggtccagat | gtcgacgcgc  | 8340 |
| ggcgccctga | ggctcagcgc | cagggtgccg | atctgagcgg | cgtccaggga | gtcgaggaag  | 8400 |
| gcctcgtgta | ggtgcagggg | cagcgtccgc | cggtggactt | gcaggagacg | ggttaagggc  | 8460 |
| ggcgccaggc | gctgatggta | cttgatctcg | agcggttcgt | tggtgagggt | gtcgatggcg  | 8520 |
| tagagcaggg | cctgaccgcg | ggcggcgacg | atggtgccgc | ggtgccggcg | gtaggtggcg  | 8580 |
| tattcggggg | ggctcgttac | atcacccgcc | tgggcctggc | gccgggcggc | agcgggggtt  | 8640 |
| ctggtccccg | cgcatgggca | ggcagcggca | cgtcggcgcg | gggctccggc | agcggctggt  | 8700 |
| gctgagctcg | cagctgactg | gcgtgcgcga | cgacgcggcg | gttgaggtcc | tggatgtgcc  | 8760 |
| tcctcgtcgt | gaagaccacc | ggtccccgga | ctcggaaact | gaaagagagt | tcgacagaat  | 8820 |
| caatctcggc | atcgttgacg | gccgcctgac | gcaggatctc | ctgcacgtcg | cccagattgt  | 8880 |
| cctggtaggc | gatctcggac | atgaactggt | cgatctcttc | ctcctggagt | tcgccgcgtc  | 8940 |
| cgggcgttcc | gacggtgccc | gcgaggtcgt | tggagatgcg | agccatgagc | tgggagaagg  | 9000 |
| cgttgaggcc | gttctcgttc | cacacgcgac | tgtagacgac | gttgccgacg | gcgtcccggg  | 9060 |
| cgcgcatgac | cacctgcgcg | acgttgagct | ccacgtgtcg | cggaagacg  | gcgtagtgtc  | 9120 |
| gcaggcgcgt | gaagaggtag | ttgaggggtg | tggcgatgtg | ctcgcagacg | aagaagtaca  | 9180 |
| tgaccacgcg | gcgacgcgtc | atctcgttga | tgtctccgag | ggcttccaag | cgctccatgg  | 9240 |
| cctcgtagaa | gtcgacggcg | aagttgaaga | actgggagtt | gcgcgccgcg | accgtcagct  | 9300 |
| cgctttgcaa | gagccggatc | agctggggca | cggtctcccg | cacctcgcgt | tcgaagggcc  | 9360 |
| ccggcgcctc | ttcctcctct | ggttcctcgg | cgccctcttc | ttccatgacg | gcttcctctt  | 9420 |
| cctccgggtc | ctcgggcacg | ggcctccggc | ggcgacggcg | cctgatgggc | aggcgggtcca | 9480 |
| cgaagcgttc | gatgatctct | ccgcggcgcg | ggcgcatggt | ttcggtgacg | gcgoggccgt  | 9540 |
| tctctcgggg | ccgcagttcg | aagacgcccc | cgcgcaggcc | gccggcgccc | ccgagagggg  | 9600 |

-continued

---

|   |       |
|---|-------|
| gcagagagtg ggggccttcg ggcagcgaga gggcgctgac gatgcaccgt atcatctgtt | 9660  |
| gcgtaggtac agctctccag gagtcgttga gcgagtcag ttggacggga tccgagaact  | 9720  |
| tttcagggaa agcttcgac caatcgacgt cgcaaggtaa gctgaggacg gtgggatgag  | 9780  |
| gggcttgggc ggaggcggag gggcgagaag aggagggaga gggcaggctg gaggtgatgc | 9840  |
| tgctgatgat gtaattgaa taggcggttt tcaaaccgag gatggtggcg aggaggacga  | 9900  |
| cgcttttggg cccggcctgc tggatgcgca ggggctggc catgccccag gcgtggctct  | 9960  |
| ggcatcggcg caggctcttg tagtagtctt gcatgagtct ctcgacgggg acgtcgtctt | 10020 |
| cgctggcccc gtcggccatg cgggtggagc cgaaccgagc caggggctgc agcaggcca  | 10080 |
| ggtcggcgac cacgccttcg gccagcacgg cctgctggat ctgggtgagc gtggcttga  | 10140 |
| agtcgtccag gtccacgaa cggtgtagg agcccgtgtt gatggtgtag gtcagttgg    | 10200 |
| ccatgacgga ccagttgacg acttgcagc cgggctgggt gatctcggtg tagcggagc   | 10260 |
| gcgagtaggc ccgcgactcg aagacgtagt cgttgcaggt gcgcacgag tactggtagc  | 10320 |
| cgacgagga gtgcgcgccg ggctcgcggt agaggggcca gcgcacggtg gcggggcg    | 10380 |
| cggggccag gtctccagc atgagcgggt gtagtggtg gacgtagcgc gagagccag     | 10440 |
| tgatgccggc ggcggaggtg gggcgcggg cgaagtcgag gacgcggttc cagatgttc   | 10500 |
| gcaagggggc gaagcgtcc atggtggca cgctctggcc ggtgagcgg gcgcagtcct    | 10560 |
| gcacgctcta gacgggacag agagcggag gttagcggct ccgctccgtg gcctggggga  | 10620 |
| cagaccgcca gggtcgacg cgggggaacc ccggttcgag accggctgga tccgtccgtc  | 10680 |
| cccgacgagc cggccccgag tccacgaccc caccagaggc cgagaccag ccgcggtgcc  | 10740 |
| cggacccag atacggagg gagccttttt gtggtttttt cccgtagatg catccggtgt   | 10800 |
| tgcgacagat gcgtccgtcg ccagcggcgc cgacgcagcc gccgctccc cccccacta   | 10860 |
| gcgcccggga ggctctgtcc ggcggccgag gcgacccgga ggaggaggcc atcctcgact | 10920 |
| tggaagaagg cgaggccctg gcccgctgg gagcgccctc ccccgagcgc catcccgcg   | 10980 |
| tgagctggc gagagactcg cggcggcct acgtgccgcc gcagaatctg ttcagggacc   | 11040 |
| gcagcggcca ggagcccag gagatgagg accgcaggtt tcacgcgggg cgggagctgc   | 11100 |
| gcgcgggctt cgaccgtcg cgggtgttc gcgccaaga cttcgagccc gacgagcga     | 11160 |
| gcggagtaag tccggcacg gcgcacgtgt cggcgccaa cctggtgacc gcgtacgagc   | 11220 |
| agacggtgaa cgaggagcgg agctttcaga aaagcttcaa caaccagtg cgcaccctga  | 11280 |
| tcgcgccga ggaggtggcc atcggcctga tgcatctgtg ggactttgtg gagcgtacg   | 11340 |
| tgcaaaccc gtcgagcaag ccgctgacgg cgcagttgtt cctgatctgt cagcacagtc  | 11400 |
| gggacaacga gacgttccg gaggcgatgc tgaacatcgc ggagcccag gccogctggc   | 11460 |
| tcttgacct gattaacatc ctgcagagca tcgtggtgca ggagcgcagc ctgagcctgg  | 11520 |
| ccgacaagg ggcggccatc aactacagca tgttgagcct gggcaagttt tacgcccga   | 11580 |
| agatctaaa gagccccctac gtgccatag acaaggagg gaagatcgac agcttttaca   | 11640 |
| tgcggatggc gctgaaagt ctgacgctga gcgacgatct gggggtgtac cgcaaccgacc | 11700 |
| gcatccaca gccctgagc gccagccgac ggcgcgagct gagcgaccgc gagctgatgc   | 11760 |
| acagcctgc gagggcgtg cggggcggc gcggcgccga ggagccgag tcctacttcg     | 11820 |
| acatgggggc ggacttgacg tggcagccca gcgcggggc cctggagcgc gcgggctacc  | 11880 |

-continued

---

gcggcgccgg cggcgtggtc gaggcggagg acgaggacga ggtggagtac gaggaggagg 11940

actgatcggc gaggtgtttt cgtagatgca gcgcgcgacg gcggcgccga gcgggcccga 12000

gggggacccc gccgtgctgg cggccctgca gagccaacct tcggcgctga acgcctccga 12060

tgactggggc gcggccatgg accgcatttt ggccttgacc acccgcaacc ccgaggcctt 12120

tagacagcag ccgcaggcca accgcttttc ggccatcttg gaagccgtgg tgccctcgcg 12180

caccaacccc acgcacgaga aggtcctggc ggtgggtaac gcgctgctgg agagcaaggc 12240

gatccgcaag gacgaggcgg gctgatttta caacgccttg ctggagcggg tggcgcgcta 12300

caacagcacc aacgtgcagg ccaacctgga ccgtctgacg acggacgtgc gggaggcggc 12360

ggcgcagcgg gagcgcttca tcgcgcgacac gaacctgggc tcgcaggtgg ccctgaacgc 12420

cttctgagc acgcagccgg ccaacgtgcc gcgcgggagc gaggactacg tcagtttcat 12480

cagcgcgctg cgccctcctg tggccgaggt gccgcagagc gaggtgtacc agtcgggtcc 12540

ggactacttc ttccagacct cgcggcaggc cctgcagacg gtgaacctga cgcaggcctt 12600

caagaacctg gaagcattgt gggcgctgcg ggccccctg ggcgaccggg cgacgatctc 12660

cagcttgctg acgccgaaca cgcgctgct gctgctgctg atcgcgcctt tcaccaatag 12720

cagtaccatc agccgcgact cgtacctggg ccacctgatc acgctgtacc gcgaggccat 12780

cgggcaggcg caggtggacg agcagacctt ccaggagatt acgagcgtga gccgggccc 12840

ggggcagcag gacacgggta gcctggaggc gacgctgaat tttctgctga ccaaccggcg 12900

gcagaagatc ccctcccagt acacgctgag cacggaggag gagcgcacct tcgcgtacgt 12960

gcagcagtc gtgagcctgt atctgatcgc cgagggggcg agcccctcgt cggcgcctgga 13020

catgacggcc cgtaacatgg agccgtcctg gtacgcggcc caccggcctg tcgtgaaccg 13080

cctgatggac tacctgcacc gcgccgccgc catgaacggc gactacttta cgaacgccat 13140

cctgaacccc cactggatgc cgcctcccg tttctacacg ggggactttg acatgcccca 13200

gggcgacgac gggttctctg gggacgacgt gtcggacagc gtgttcgcgc cggtgctcc 13260

gggcaagaag gagggcggcg acgagctgcc gctgtccgtg gtggaggcgg cgtcgcgcgg 13320

ccagagcccc ttcccagcc tcccgtcgtt gtcggcagc agcagcagcg gccgggtctc 13380

gcgcccgcgg ctggaggcgg actacctgaa cgacccgctg ctgcgcgccg cccggcccga 13440

gaactttccc aacaaccggg tggagagcct agtggataag atgaatcgtt ggaagacct 13500

cgcccaggag cagcgggagt gggaggagag tcagccccgc cccctgcctc cgcgcgcctc 13560

caggtggcgc cgcggggaag aagaccggga agactcggcg gacgatagca gcgtgttgg 13620

cttggggggg accggtgcc cctcgacaaa cccgttgc caccctgcgc cgcagggccg 13680

gctgggtcgg ctgtattgag gaaagaaact aataaaagaa aaaagactt gcttaccaga 13740

gccatggtcg cagcgtcggc ccctttgtgt gtgttttctc ctccccgta gcgaa atg  
Met  
1

agg cgc gcg gtg gga gtg ccg ccg gtg atg gcg tac gcc gag ggt cct 13846  
Arg Arg Ala Val Gly Val Pro Pro Val Met Ala Tyr Ala Glu Gly Pro  
5 10 15

cct cct tet tac gaa acg gtg atg ggc gcc gcg gat tcg ccg gcc acg 13894  
Pro Pro Ser Tyr Glu Thr Val Met Gly Ala Ala Asp Ser Pro Ala Thr  
20 25 30

ctg gag gcg ctc tac gtc cct ccc cgc tac ctg ggg cct acg gag ggg 13942  
Leu Glu Ala Leu Tyr Val Pro Pro Arg Tyr Leu Gly Pro Thr Glu Gly



-continued

| 35  | 40 | 45 |       |
|---|----|----|-------|
| agg aac agc atc cgt tac tca gag ctg gcg ccg ctg tac gac acc acc<br>Arg Asn Ser Ile Arg Tyr Ser Glu Leu Ala Pro Leu Tyr Asp Thr Thr<br>50 55 60 65     |    |    | 13990 |
| cgc gtg tac ctg gtg gat aac aag tcg gcg gac atc gcg tcg ctg aac<br>Arg Val Tyr Leu Val Asp Asn Lys Ser Ala Asp Ile Ala Ser Leu Asn<br>70 75 80        |    |    | 14038 |
| tac cag aac gac cat agc aac ttt ctg acc acg gtg gtg cag aac aat<br>Tyr Gln Asn Asp His Ser Asn Phe Leu Thr Thr Val Val Gln Asn Asn<br>85 90 95        |    |    | 14086 |
| gac ttt acc ccg gtg gag gcg ggc acg cag acc ata aat ttc gac gag<br>Asp Phe Thr Pro Val Glu Ala Gly Thr Gln Thr Ile Asn Phe Asp Glu<br>100 105 110     |    |    | 14134 |
| cgc tcg cgg tgg ggc ggc gac ctg aaa acc atc ctg cgc acc aac atg<br>Arg Ser Arg Trp Gly Gly Asp Leu Lys Thr Ile Leu Arg Thr Asn Met<br>115 120 125     |    |    | 14182 |
| ccc aac atc aac gag ttc atg tcc acc aac aag ttc agg gcc cgg ttg<br>Pro Asn Ile Asn Glu Phe Met Ser Thr Asn Lys Phe Arg Ala Arg Leu<br>130 135 140 145 |    |    | 14230 |
| atg gta gag aaa gtg aac aag gaa acc aat gcc cct cga tac gag tgg<br>Met Val Glu Lys Val Asn Lys Glu Thr Asn Ala Pro Arg Tyr Glu Trp<br>150 155 160     |    |    | 14278 |
| ttt gag ttc acc ctg ccc gag ggc aac tac tcg gag acc atg acc ata<br>Phe Glu Phe Thr Leu Pro Glu Gly Asn Tyr Ser Glu Thr Met Thr Ile<br>165 170 175     |    |    | 14326 |
| gac ctg atg aat aac gcg atc gtg gac aac tac ttg gaa gtg ggg cgg<br>Asp Leu Met Asn Asn Ala Ile Val Asp Asn Tyr Leu Glu Val Gly Arg<br>180 185 190     |    |    | 14374 |
| cag aac ggg gtg ctg gag agc gac atc ggg gtg aag ttt gac acg cgc<br>Gln Asn Gly Val Leu Glu Ser Asp Ile Gly Val Lys Phe Asp Thr Arg<br>195 200 205     |    |    | 14422 |
| aac ttc cgg ctg ggc tgg gac ccg gtc acc aag ctg gtc atg ccc ggc<br>Asn Phe Arg Leu Gly Trp Asp Pro Val Thr Lys Leu Val Met Pro Gly<br>210 215 220 225 |    |    | 14470 |
| gtg tac acc aac gag gcc ttc cac ccc gac atc gtc ctg ctg ccc ggc<br>Val Tyr Thr Asn Glu Ala Phe His Pro Asp Ile Val Leu Leu Pro Gly<br>230 235 240     |    |    | 14518 |
| tgc ggc gtg gac ttc acg cag agc cgg ctg agc aac ctg ctg ggg atc<br>Cys Gly Val Asp Phe Thr Gln Ser Arg Leu Ser Asn Leu Leu Gly Ile<br>245 250 255     |    |    | 14566 |
| cgc aag cgg atg ccc ttc cag gcg ggt ttt cag atc atg tac gag gac<br>Arg Lys Arg Met Pro Phe Gln Ala Gly Phe Gln Ile Met Tyr Glu Asp<br>260 265 270     |    |    | 14614 |
| ctg gag ggc ggc aac atc ccc gcc ttg cta gac gtg gcg aaa tac gag<br>Leu Glu Gly Gly Asn Ile Pro Ala Leu Leu Asp Val Ala Lys Tyr Glu<br>275 280 285     |    |    | 14662 |
| gcc agc att cag aag gcg cgg gag cag gcc cag gag atc cgc ggc gac<br>Ala Ser Ile Gln Lys Ala Arg Glu Gln Gly Gln Glu Ile Arg Gly Asp<br>290 295 300 305 |    |    | 14710 |
| aac ttt acc gtc atc ccc cgg gac gtg gag atc gtg ccc gtg gag aag<br>Asn Phe Thr Val Ile Pro Arg Asp Val Glu Ile Val Pro Val Glu Lys<br>310 315 320     |    |    | 14758 |
| gat agc aag gac cgc agt tac aac cta ctc ccc gcc gac cag acc aac<br>Asp Ser Lys Asp Arg Ser Tyr Asn Leu Leu Pro Gly Asp Gln Thr Asn<br>325 330 335     |    |    | 14806 |
| acg gcc tac cgc agc tgg ttc ctg gcc tac aac tac gcc gac ccc gag<br>Thr Ala Tyr Arg Ser Trp Phe Leu Ala Tyr Asn Tyr Gly Asp Pro Glu<br>340 345 350     |    |    | 14854 |

-continued

| 340  | 345 | 350 |       |
|--|-----|-----|-------|
| aag ggc gtc agg tcc tgg acg ctg ctg acc acc acg gac gtc acc tgc    |     |     | 14902 |
| Lys Gly Val Arg Ser Trp Thr Leu Leu Thr Thr Thr Asp Val Thr Cys    |     |     |       |
| 355  | 360 | 365 |       |
| ggc tcg cag cag gtg tac tgg tcg ctc ccg gac atg atg caa gac ccc    |     |     | 14950 |
| Gly Ser Gln Gln Val Tyr Trp Ser Leu Pro Asp Met Met Gln Asp Pro    |     |     |       |
| 370  | 375 | 380 | 385   |
| gtg acc ttc cgg ccc tcc agc caa gtc agc aac tac ccc gtg gtg gga    |     |     | 14998 |
| Val Thr Phe Arg Pro Ser Ser Gln Val Ser Asn Tyr Pro Val Val Gly    |     |     |       |
|  | 390 | 395 | 400   |
| gtc gag ctc ctg ccg gtg cac gcc aag agc ttt tac aac gag cag gcc    |     |     | 15046 |
| Val Glu Leu Leu Pro Val His Ala Lys Ser Phe Tyr Asn Glu Gln Ala    |     |     |       |
|  | 405 | 410 | 415   |
| gtc tac tcg cag ctc atc cgc cag tcc acc gcg ctc acg cac gtc ttc    |     |     | 15094 |
| Val Tyr Ser Gln Leu Ile Arg Gln Ser Thr Ala Leu Thr His Val Phe    |     |     |       |
|  | 420 | 425 | 430   |
| aac cgc ttc ccc gag aac cag atc ctg gtg cgc ccg ccc gct ccg acc    |     |     | 15142 |
| Asn Arg Phe Pro Glu Asn Gln Ile Leu Val Arg Pro Pro Ala Pro Thr    |     |     |       |
|  | 435 | 440 | 445   |
| att acc acc gtc agt gaa aac gtt ccc gcc ctc aca gat cac gga acc    |     |     | 15190 |
| Ile Thr Thr Val Ser Glu Asn Val Pro Ala Leu Thr Asp His Gly Thr    |     |     |       |
|  | 450 | 455 | 460   |
| ctg ccg ctg cgc agc agt atc agt gga gtc cag cgc gtg acc atc act    |     |     | 15238 |
| Leu Pro Leu Arg Ser Ser Ile Ser Gly Val Gln Arg Val Thr Ile Thr    |     |     |       |
|  | 470 | 475 | 480   |
| gac gcc cgg cga agg acc tgc ccc tac gtg cac aag gcc ctg ggc ata    |     |     | 15286 |
| Asp Ala Arg Arg Arg Thr Cys Pro Tyr Val His Lys Ala Leu Gly Ile    |     |     |       |
|  | 485 | 490 | 495   |
| gtc gct ccc aaa gtg ctc tct agc cgc acc ttt taa caagcatgtc         |     |     | 15332 |
| Val Ala Pro Lys Val Leu Ser Ser Arg Thr Phe                        |     |     |       |
|  | 500 | 505 |       |
| cattctcctc tcgcccgaca acaacaccgg ctggggcctg cgctcggccg gcatgtacgg  |     |     | 15392 |
| cgggcccaag cggcgctcca gcgagcacc ccgtccgcgtc cggggccact accgggcccc  |     |     | 15452 |
| ctggggcgcc cacaagcgcg gcgtctccac gcgcaccacc gtcgacgacg ccatcgacgc  |     |     | 15512 |
| cgctgtggcc caggccagac gctaccggccg gcccaagtgc acggtggacg ccgtcatcga |     |     | 15572 |
| cagcgtggty gccgacgcgc ggcgatacgc tcgacgcaag cggcgtctgc accgccgtcg  |     |     | 15632 |
| ccgtcccacc gccgcatgc tggccgccag agcggctctg agacgcgcgc gccgcgtggg   |     |     | 15692 |
| ccgcgagacc atgcgccgag ccgcggccaa cgcacgcgcg ggtcgcgccc gtcgtcaggc  |     |     | 15752 |
| cgcccggcag gccgcccgcg ccatcgccaa cctggcccaa ccccgcggg gaaacgtgta   |     |     | 15812 |
| ctgggtgcga gacgcgtcgg gcgtgcgcgt gccggtgcgc acccgcccc ctcggagtta   |     |     | 15872 |
| gaagacaaaa agacggacga agactgagtt tcctgtcgt tgccagcatg agcaagcgca   |     |     | 15932 |
| agttcaaaga agagctgctg gaggccctcg tgcccagat ctacggcccg gccgcccgtg   |     |     | 15992 |
| ccgcgcgggt gccggacgtc aagcccgaag ttaagcccc cgcgctgaag cgggttaaaa   |     |     | 16052 |
| agcgggaaaa gaaagaggag aagcaggaag cagggttgct agacgtcgac gacggcgtgg  |     |     | 16112 |
| agttcgtcgc gtccttcgcg ccccgtcgcc gggtcagtg gcggggtcgc cgcgtcaagc   |     |     | 16172 |
| tcgtcccgcg gccgggcacc gtggtgtctt tcacccccg cctgcgttcg gccacgcgcg   |     |     | 16232 |
| gcctgaagcg cgagtacgac gaggctctat gcgacgaaga catcctggag caggccgccc  |     |     | 16292 |
| agcagctcgg ggagtttgct tacggcaagc gcggccgcta cggggaggtg gcgctggcgc  |     |     | 16352 |

-continued

---

|  |                 |
|--|-----------------|
| tggaccaggg caatcccacg cccagcctca agcccgtcac gctgcagcag gtgctgcccg  | 16412           |
| tgagcgcgtc gaccgagagc aagcggggca tcaagaggga gatgggcgac ctgcagccca  | 16472           |
| ccatgcaact catggtgccc aaacggcaga agctggagga cgtgctggag aacatgaaag  | 16532           |
| tggatcccag catcgagccc gaagtgaaag tgcgacccat caaggaagtg ggcccgggccc | 16592           |
| taggcgtgca gacggtggac attcagatcc ccgtgcgcgc ctccccggtt tctgcccacca | 16652           |
| ctacgacggc cgtggaggcc atggaaacgc agacggagct gcccgcggcc ttggcggcag  | 16712           |
| ccgccaccgc cggcgggct acccgagaga tgggcatgca gaccgacccc tggtagcagt   | 16772           |
| tcgcccggccc cggccgtcgt ccacgagccc gtcggtacgc ggcgaccacc tcccggctcc | 16832           |
| ctgactacgt cttgcatcct tccatcacgc cgacgcccgg ctaccgggga acgaccttc   | 16892           |
| gccccggctg cgcgcgcacc accacccgcc gtcgtcgcac caccggccgc cgtgcgagcc  | 16952           |
| gtcgcgcact ggctcccacg gcggttcgcc gcgtcgtccg ccggggctgc acgctgacct  | 17012           |
| tgcccaccgc gcgttaccac cccagcatcg tcatttaacc tgcgctgccg ttttgagat   | 17072           |
| ggctctgacg tgcgctttc gcttccccgt tcggcactac cgaggaagat ctgcccgtag   | 17132           |
| gactgttcta gcgggcagcg gtctccgacg ccgcccgcgc gcggtgcacc ggcgcatgaa  | 17192           |
| ggcgccatt ctgcccgcgc tgatccccat tatcgccgcc gccatcgggg cgatccccgg   | 17252           |
| cgtggcctcg gtggccttgc aagcagctcg caaaaattaa ataagaagg cttgacactc   | 17312           |
| actgcctggt cctgactggt tcatgcagac aagacatgga agacatcaat tttgcgtcgt  | 17372           |
| tggccccgcg gcacggctcg cggccgttca tgggcacctg gaacgagatc ggcaccagcc  | 17432           |
| agctcaacgg gggcgctttc agttggagca gcctgtggag cggcattaaa aactttgggt  | 17492           |
| ccacgattaa gacctatgac aacaaggcgt ggaacagtag cactggtcag atgctccgcg  | 17552           |
| ataagctgaa ggaccagaac ttccagcaga aagtggtaga cgtctggcc tcgggcatca   | 17612           |
| acgggggtgt ggacctggcc aaccaggcgg tgcagaacca gatcaaccag cgtctggaga  | 17672           |
| acagccgcca gccgcccgcg gccctgcagc agcgtccgca ggtggaggag gtggaagtgg  | 17732           |
| aggagaagct gccgcccctg gagacggtgt cgccgggtgg cgtgcctagc aagggggaga  | 17792           |
| agcggccgcg gcccgagctc gaggagacc tagtgaccga gaccctggag ccccccctgt   | 17852           |
| acgagcaggc cttgaaagag ggggcccacg ccctgcccac gaccggccc atcggacca    | 17912           |
| tggcccagcc ggtctacgac aaggaacaca aagccgtgac gctagagctg cctccgcccg  | 17972           |
| cgcccaccgt acccccgatg cccggtccca cctggggcac gcgctgctc cgtcccgcg    | 18032           |
| ccccgccggt cgcggtggcc acgcccgcgc gcccgagtgc cggagccaac tggcagagca  | 18092           |
| ctctgaacag catcgtgggc ctgggagtga aaagcctgaa acgcccggg tgttactatt   | 18152           |
| aaagccagct aaatacccat gtgtgtatg cgcctcctgt gtcacgccag aaaaagccag   | 18212           |
| ccgagtgacg ggtcaccgcc gccgccaaga gcgcccgttt caag atg gcc acc ccc   | 18268           |
|  | Met Ala Thr Pro |
|  | 510             |
| tcg atg atg ccg cag tgg tct tac atg cac atc gcc ggg cag gac gcc    | 18316           |
| Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ala Gly Gln Asp Ala    |                 |
| 515 520 525  |                 |
| tcg gag tac ctg agc ccg ggc ctg gtg cag ttc gcc cgc gcc acc gac    | 18364           |
| Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala Arg Ala Thr Asp    |                 |
| 530 535 540  |                 |
| acg tac ttc agc ctg ggc aac aag ttt agg aac ccc acg gtg gcc ccc    | 18412           |
| Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro Thr Val Ala Pro    |                 |

-continued

| 545   | 550 | 555 | 560 |       |
|---|-----|-----|-----|-------|
| acc cac gac gtg acg acg gac cgg tcc cag cgg ctg acg ctg cgg ttc |     |     |     | 18460 |
| Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu Thr Leu Arg Phe | 565 | 570 | 575 |       |
| gtg ccc gtc gac cgc gag gac acc cgg tac tcg tac aaa gtg cgc ttc |     |     |     | 18508 |
| Val Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr Lys Val Arg Phe | 580 | 585 | 590 |       |
| acg ctg gcc gtg ggc gac aac cgc gtg ctg gac atg gcc agc acg tac |     |     |     | 18556 |
| Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met Ala Ser Thr Tyr | 595 | 600 | 605 |       |
| ttt gac atc cgc ggc gtg ttg gac cgc ggt ccc agc ttc aaa ccc tac |     |     |     | 18604 |
| Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser Phe Lys Pro Tyr | 610 | 615 | 620 |       |
| tcc gcc acc gcc tac aac tcc ctg gcc ccc aag ggc gcc ccc aac ccg |     |     |     | 18652 |
| Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly Ala Pro Asn Pro | 625 | 630 | 635 | 640   |
| tca gaa tgg aag ggc tca gac aac aaa att agt gta aga ggt cag gct |     |     |     | 18700 |
| Ser Glu Trp Lys Gly Ser Asp Asn Lys Ile Ser Val Arg Gly Gln Ala | 645 | 650 | 655 |       |
| ccg ttt ttt agt aca tcc att aca aag gat ggt att caa gtg gcc act |     |     |     | 18748 |
| Pro Phe Phe Ser Thr Ser Ile Thr Lys Asp Gly Ile Gln Val Ala Thr | 660 | 665 | 670 |       |
| gat act tct agc gga gct gtg tat gct aaa aag gaa tat cag cct gaa |     |     |     | 18796 |
| Asp Thr Ser Ser Gly Ala Val Tyr Ala Lys Lys Glu Tyr Gln Pro Glu | 675 | 680 | 685 |       |
| cca caa gta ggg caa gaa caa tgg aac agc gaa gcc agt gat agt gat |     |     |     | 18844 |
| Pro Gln Val Gly Gln Glu Gln Trp Asn Ser Glu Ala Ser Asp Ser Asp | 690 | 695 | 700 |       |
| aaa gta gct ggt agg att cta aaa gac aca aca ccc atg ttc cct tgt |     |     |     | 18892 |
| Lys Val Ala Gly Arg Ile Leu Lys Asp Thr Thr Pro Met Phe Pro Cys | 705 | 710 | 715 | 720   |
| tac ggt tcc tac gcc aag ccc aca aat gaa cag ggg ggg caa ggc act |     |     |     | 18940 |
| Tyr Gly Ser Tyr Ala Lys Pro Thr Asn Glu Gln Gly Gly Gln Gly Thr | 725 | 730 | 735 |       |
| aat act gta gat ctg cag ttc ttt gcc tct tca tcg gct acc tct acg |     |     |     | 18988 |
| Asn Thr Val Asp Leu Gln Phe Phe Ala Ser Ser Ser Ala Thr Ser Thr | 740 | 745 | 750 |       |
| cct aaa gcc gta ctc tat gcc gag gac gtg gca ata gaa gca cca gac |     |     |     | 19036 |
| Pro Lys Ala Val Leu Tyr Ala Glu Asp Val Ala Ile Glu Ala Pro Asp | 755 | 760 | 765 |       |
| acc cat ttg gtg tac aaa ccg gca gtt aca acc acg acc act agt tcc |     |     |     | 19084 |
| Thr His Leu Val Tyr Lys Pro Ala Val Thr Thr Thr Thr Ser Ser     | 770 | 775 | 780 |       |
| caa gac ctg cta act cag cag gct gct ccc aac cga ccc aac tac att |     |     |     | 19132 |
| Gln Asp Leu Leu Thr Gln Gln Ala Ala Pro Asn Arg Pro Asn Tyr Ile | 785 | 790 | 795 | 800   |
| ggc ttc agg gat aat ttt atc ggt ctc atg tat tac aac tcc act ggc |     |     |     | 19180 |
| Gly Phe Arg Asp Asn Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly | 805 | 810 | 815 |       |
| aat atg ggt gtt ttg gca ggg caa gct tct cag cta aac cgc gtg gtt |     |     |     | 19228 |
| Asn Met Gly Val Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val | 820 | 825 | 830 |       |
| gac ttg caa gac aga aac acc gag ctg tcc tac cag ctc atg ctt gat |     |     |     | 19276 |
| Asp Leu Gln Asp Arg Asn Thr Glu Leu Ser Tyr Gln Leu Met Leu Asp | 835 | 840 | 845 |       |
| gct ttg ggc gac cgc agt cgt tac ttc tcc atg tgg aac cag gcc gta |     |     |     | 19324 |
| Ala Leu Gly Asp Arg Ser Arg Tyr Phe Ser Met Trp Asn Gln Ala Val |     |     |     |       |

-continued

| 850         | 855                         | 860                             |       |
|-------------|-----------------------------|---------------------------------|-------|
| gac agc tat | gac cct gat gtc aga att att | gaa aat cat ggt gtg gag         | 19372 |
| Asp Ser Tyr | Asp Pro Asp Val Arg Ile Ile | Glu Asn His Gly Val Glu         |       |
| 865         | 870                         | 875 880                         |       |
| gat gag ctg | cca aac tac tgt ttc ccg cta | gga ggg tcg cta gta act         | 19420 |
| Asp Glu Leu | Pro Asn Tyr Cys Phe Pro     | Leu Gly Gly Ser Leu Val Thr     |       |
|             | 885                         | 890 895                         |       |
| gaa act tat | aca ggc cta tca ccc caa aac | gga agt aac acg tgg aca         | 19468 |
| Glu Thr Tyr | Thr Gly Leu Ser Pro Gln     | Asn Gly Ser Asn Thr Trp Thr     |       |
|             | 900                         | 905 910                         |       |
| acc gac agc | acc acc tat gca act aga ggg | gtg gaa atc ggc tct ggc         | 19516 |
| Thr Asp Ser | Thr Thr Tyr Ala Thr Arg     | Gly Val Glu Ile Gly Ser Gly     |       |
|             | 915                         | 920 925                         |       |
| aac atg ttc | gcc atg gaa att aat ttg gcg | gcc aat cta tgg agg agt         | 19564 |
| Asn Met Phe | Ala Met Glu Ile Asn Leu     | Ala Ala Asn Leu Trp Arg Ser     |       |
|             | 930                         | 935 940                         |       |
| ttc ctg tac | tcc aac gtg gcc ctg tac ctg | ccc gac gag tac aag ctc         | 19612 |
| Phe Leu Tyr | Ser Asn Val Ala Leu Tyr     | Leu Pro Asp Glu Tyr Lys Leu     |       |
|             | 945                         | 950 955 960                     |       |
| acc ccc gac | aac atc acc ctc ccc gac aac | aaa aac act tac gac tac         | 19660 |
| Thr Pro Asp | Asn Ile Thr Leu Pro Asp     | Asn Lys Asn Thr Tyr Asp Tyr     |       |
|             | 965                         | 970 975                         |       |
| atg aac ggc | cgc gtg gcc gcc ccc agc tcc | ctc gac acc tac gtc aac         | 19708 |
| Met Asn Gly | Arg Val Ala Ala Pro Ser     | Ser Ser Leu Asp Thr Tyr Val Asn |       |
|             | 980                         | 985 990                         |       |
| atc ggg gcg | cgc tgg tcc ccc gac ccc     | atg gac aac gtc aac ccc ttc     | 19756 |
| Ile Gly Ala | Arg Trp Ser Pro Asp Pro     | Met Asp Asn Val Asn Pro Phe     |       |
|             | 995                         | 1000 1005                       |       |
| aac cac cac | cgc aac gcg gga ctg cgc     | tac cgc tcc atg ctg ctg         | 19801 |
| Asn His His | Arg Asn Ala Gly Leu Arg     | Tyr Arg Ser Met Leu Leu         |       |
|             | 1010                        | 1015 1020                       |       |
| ggc aac ggc | cgc tac gta ccc ttc cac     | atc caa gtg ccc cag aaa         | 19846 |
| Gly Asn Gly | Arg Tyr Val Pro Phe         | His Ile Gln Val Pro Gln Lys     |       |
|             | 1025                        | 1030 1035                       |       |
| ttc ttc gcc | atc aaa aac ctc ctg ctc     | ctc ccc ggg tcc tac acc         | 19891 |
| Phe Phe Ala | Ile Lys Asn Leu Leu Leu     | Leu Pro Gly Ser Tyr Thr         |       |
|             | 1040                        | 1045 1050                       |       |
| tac gag tgg | aac ttc cgc aag gac gtc     | aac atg atc ctc cag agc         | 19936 |
| Tyr Glu Trp | Asn Phe Arg Lys Asp         | Val Asn Met Ile Leu Gln Ser     |       |
|             | 1055                        | 1060 1065                       |       |
| agc ctg ggt | aac gac ctc cgc gtc gac     | ggg gcc agc gtc agg ttc         | 19981 |
| Ser Leu Gly | Asn Asp Leu Arg Val         | Asp Gly Ala Ser Val Arg Phe     |       |
|             | 1070                        | 1075 1080                       |       |
| gac agc atc | aac ctg tac gcc aac ttc     | ttc ccc atg gcc cac aac         | 20026 |
| Asp Ser Ile | Asn Leu Tyr Ala Asn         | Phe Phe Pro Met Ala His Asn     |       |
|             | 1085                        | 1090 1095                       |       |
| acc gcc tcc | acg ctc gag gcc atg ctg     | cgc aac gac acc aac gac         | 20071 |
| Thr Ala Ser | Thr Leu Glu Ala Met         | Leu Arg Asn Asp Thr Asn Asp     |       |
|             | 1100                        | 1105 1110                       |       |
| cag tcg ttc | aac gac tac ctc tgc gct     | gcc aac atg ctc tac ccc         | 20116 |
| Gln Ser Phe | Asn Asp Tyr Leu Cys         | Ala Ala Asn Met Leu Tyr Pro     |       |
|             | 1115                        | 1120 1125                       |       |
| atc ccc gcc | aac gcc acc agc gtg ccc     | atc tcc att ccc tcg cg          | 20161 |
| Ile Pro Ala | Asn Ala Thr Ser Val         | Pro Ile Ser Ile Pro Ser Arg     |       |
|             | 1130                        | 1135 1140                       |       |
| aac tgg gcc | gcc ttc cg                  | ggc tgg agc ttc acc             | 20206 |
| Asn Trp Ala | Ala Phe Arg Gly             | Trp Ser Phe Thr Arg Leu Lys Thr |       |

-continued

|   |      |      |  |
|---|------|------|--|
| 1145  | 1150 | 1155 |  |
| aag gag acc ccc tct ctg ggc tcc ggc ttc gat ccc tac ttc acc 20251       |      |      |  |
| Lys Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr Phe Thr             |      |      |  |
| 1160  | 1165 | 1170 |  |
| tac tcg ggc tcc atc ccc tac ctg gac ggc acc ttc tac ctc aac 20296       |      |      |  |
| Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu Asn             |      |      |  |
| 1175  | 1180 | 1185 |  |
| cac act ttc aag aag gtc tcc atc atg ttc gac tcc tcc gtc agc 20341       |      |      |  |
| His Thr Phe Lys Lys Val Ser Ile Met Phe Asp Ser Ser Val Ser             |      |      |  |
| 1190  | 1195 | 1200 |  |
| tgg ccc ggc aac gac cgc ctg ctg acc ccc aac gag ttc gag atc 20386       |      |      |  |
| Trp Pro Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu Ile             |      |      |  |
| 1205  | 1210 | 1215 |  |
| aag cgc acc gtg gac ggg gaa ggg tac aac gtg gcc cag tgc aac 20431       |      |      |  |
| Lys Arg Thr Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn             |      |      |  |
| 1220  | 1225 | 1230 |  |
| atg acc aag gac tgg ttc ctc atc cag atg ctc agc cac tac aac 20476       |      |      |  |
| Met Thr Lys Asp Trp Phe Leu Ile Gln Met Leu Ser His Tyr Asn             |      |      |  |
| 1235  | 1240 | 1245 |  |
| atc ggc tac cag ggc ttc tac gtg ccc gag ggc tac aag gac agg 20521       |      |      |  |
| Ile Gly Tyr Gln Gly Phe Tyr Val Pro Glu Gly Tyr Lys Asp Arg             |      |      |  |
| 1250  | 1255 | 1260 |  |
| atg tac tct ttc ttc cgc aac ttc caa ccc atg agc cgc cag gtg 20566       |      |      |  |
| Met Tyr Ser Phe Phe Arg Asn Phe Gln Pro Met Ser Arg Gln Val             |      |      |  |
| 1265  | 1270 | 1275 |  |
| gtc gac acc acc acc tac acc gac tac aaa aac gtc acc ctc ccc 20611       |      |      |  |
| Val Asp Thr Thr Thr Tyr Thr Asp Tyr Lys Asn Val Thr Leu Pro             |      |      |  |
| 1280  | 1285 | 1290 |  |
| ttc cag cac aac aac tcg ggg ttc gtg gga tac atg ggc ccc acc 20656       |      |      |  |
| Phe Gln His Asn Asn Ser Gly Phe Val Gly Tyr Met Gly Pro Thr             |      |      |  |
| 1295  | 1300 | 1305 |  |
| atg cgc gag ggg cag gcc tac ccc gcc aac tac ccc tac ccc ctg 20701       |      |      |  |
| Met Arg Glu Gly Gln Ala Tyr Pro Ala Asn Tyr Pro Tyr Pro Leu             |      |      |  |
| 1310  | 1315 | 1320 |  |
| atc ggc aag acc gcc gtg ccc agc ctc acg cag aaa aag ttc ctc 20746       |      |      |  |
| Ile Gly Lys Thr Ala Val Pro Ser Leu Thr Gln Lys Lys Phe Leu             |      |      |  |
| 1325  | 1330 | 1335 |  |
| tgc gac cgc acc atg tgg cgc atc ccc ttc tcc agt aac ttc atg 20791       |      |      |  |
| Cys Asp Arg Thr Met Trp Arg Ile Pro Phe Ser Ser Asn Phe Met             |      |      |  |
| 1340  | 1345 | 1350 |  |
| tcc atg ggg gcg ctc acc gac ctg ggg cag aac atg ctg tac gcc 20836       |      |      |  |
| Ser Met Gly Ala Leu Thr Asp Leu Gly Gln Asn Met Leu Tyr Ala             |      |      |  |
| 1355  | 1360 | 1365 |  |
| aac tcc gcc cac gcc ctc gac atg acc ttc gag gtg gac ccc atg 20881       |      |      |  |
| Asn Ser Ala His Ala Leu Asp Met Thr Phe Glu Val Asp Pro Met             |      |      |  |
| 1370  | 1375 | 1380 |  |
| gat gag ccc acg ctt ctc tat gtt ctg ttc gaa gtg ttc gac gtc 20926       |      |      |  |
| Asp Glu Pro Thr Leu Leu Tyr Val Leu Phe Glu Val Phe Asp Val             |      |      |  |
| 1385  | 1390 | 1395 |  |
| gtg cgc atc cac cag ccg cac cgc ggc gtc atc gag gcc gtc tac 20971       |      |      |  |
| Val Arg Ile His Gln Pro His Arg Gly Val Ile Glu Ala Val Tyr             |      |      |  |
| 1400  | 1405 | 1410 |  |
| ctg cgc acg ccg ttc tcg gcc ggt aac gcc acc acc taa ggagggggcc 21020    |      |      |  |
| Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr                         |      |      |  |
| 1415  | 1420 | 1425 |  |
| gccgacggat gggctccagc gagccggagc tggctgccat cgcgcgcgac ctgggctgcg 21080 |      |      |  |

-continued

---

|            |            |            |            |            |             |       |
|------------|------------|------------|------------|------------|-------------|-------|
| ggccctactt | cctgggcacc | tttgacaaac | gcttcccggg | cttcgtggcg | ccgcacaagc  | 21140 |
| tggcctgctc | catcgtcaac | accgcccggc | gagagaccgg | cgccgtccac | tgctggccc   | 21200 |
| tgccctggaa | cccccgagc  | cgaacctgct | acctcttcca | ccccttcggc | ttctcggagc  | 21260 |
| acaggtctag | gcagatctac | cagttcaggt | acgaaggcct | gctccggcgc | agcgcctctg  | 21320 |
| cctccacccc | cgaccactgc | gtcacctctg | tcaagtccac | ccagaccgtc | caggggcccc  | 21380 |
| gctcggccgc | ctgcccctc  | ttctgctgca | tgctcctgca | cgctctctg  | cgctggccc   | 21440 |
| cctcccccat | ggagcgcaac | cccaccatgg | acctccttac | ggcggttccc | aacagcatgc  | 21500 |
| ttcagagtcc | ccaggtcag  | cccacctcc  | accgcaacca | ggaggaactc | tacgccttcc  | 21560 |
| tggtcggca  | ctccccctac | tttcgccc   | accgagagcg | catagaaaag | gccaccgcgt  | 21620 |
| ttgacaaaat | gaacgactag | atctctctg  | aaaaactc   | aataaagcct | ttattggttc  | 21680 |
| accacacgtg | cagcatgca  | gactttttat | ttaaagggc  | tccgctcct  | cgctcggctg  | 21740 |
| gctggtggg  | agggagagct | tcgatactg  | cagggggag  | ctccatctga | actcgggaat  | 21800 |
| cagcagcttg | ggcagggggc | cctcagctt  | ctcgtccac  | agcttgca   | ccagctgag   | 21860 |
| ggcggccagc | aggtcggg   | cggagatctt | gaagtcgag  | ttggggcct  | ggttgcggc   | 21920 |
| ggagttgctg | tacaccgggt | tgccgactg  | gaacaccagc | acgtgggg   | gctcagctg   | 21980 |
| ggccagcgcc | gtcttgctg  | tcacctctg  | gcccgcagg  | gactccgct  | tgctcagcg   | 22040 |
| gaaggggctc | agcttgca   | gtgcccagc  | cagcacggc  | acccgctc   | gctggttcag  | 22100 |
| gcagtcgag  | cgcatagcca | tcagcagcg  | cttctgccc  | tgctgcatct | tcggatagtc  | 22160 |
| ggctcagctg | aaggcctcca | ctgcccga   | ggcgtctg   | gccttgctg  | cctccgagaa  | 22220 |
| gaacagccc  | caggacttg  | cggagaacac | gttgttgcg  | cagctcagct | ctccacgca   | 22280 |
| gcagcggc   | tcgtcttct  | tcagctgca  | cacgtcgg   | cccagcgg   | ctgcaaccac  | 22340 |
| cttggctctg | ccgggatgt  | cctcagggc  | ccgtggcc   | ttctcgttg  | tcagctccat  | 22400 |
| ctccaccacc | tgctcctct  | ggatcatctc | cagcccgtg  | tagcagcga  | gcacgcctc   | 22460 |
| ctgctcgg   | caccctgca  | gccagacgc  | gcagccgtc  | ggctccagct | ggtgaggtt   | 22520 |
| caccccgcg  | taggtctcca | cgtacccc   | caggaagcg  | ccatcatct  | ccacaaaggt  | 22580 |
| ctctgaccg  | gtgaaggta  | gtgccc     | gcgatgctc  | tcgttgagc  | acgtctgaca  | 22640 |
| gatcttgcg  | tacacctg   | cctgctcgg  | cagaaactg  | aaagcggcct | tctcctcgg   | 22700 |
| ctccacgtg  | tacttctcca | tcagcggca  | catcagctc  | atgctctct  | cccagccga   | 22760 |
| caccagcgg  | tccgcccgg  | ggttcaccac | cgccatgct  | cgggaagtgc | cggggcgctc  | 22820 |
| atcttctct  | tcctcctct  | cttctcttg  | aggcggcgg  | ggcggcagtt | gtctcagaa   | 22880 |
| tctcttgcc  | ttggccttct | ggacgatctc | caccccggg  | tggtgaacc  | cgctggccac  | 22940 |
| caccacttcg | tcctctctct | cttcgctg   | ggcacgact  | tcgggagag  | gagcggcgg   | 23000 |
| aggaaccgg  | gcggccactg | cggccatc   | ggcgttctg  | cgcgctctct | tgggggcag   | 23060 |
| aggcggcgtc | tcgctctcc  | ggctgctc   | ttgcagtag  | ggcgtgatg  | tgtgggaggt  | 23120 |
| ggggcgtct  | ggctgacgc  | cggccatgct | gatgcttgac | tcctagcga  | aaagatggag  | 23180 |
| gaggatctta | gacagccga  | gcccgtctc  | gaaacctta  | ccaccccg   | ctctgaggtc  | 23240 |
| ggcggcggc  | agctagacat | gcaacgggag | gaggaggag  | acgtcagag  | ggagcaagac  | 23300 |
| ccgggctacg | tgacccc    | cgaggacgc  | gaggagccg  | aggaccggc  | gccaaccgctc | 23360 |

-continued

---

|  |       |
|--|-------|
| agcgaagccg actacctggg aggggaggac gacgtgctgc tgaagcacct ggcgcggcag  | 23420 |
| agcaccatcg tgcaggaggc cctcaaggag cgcgaggagg tcccgcctgac ggtggaggag | 23480 |
| ctcagccggg cctacgaagc caacctcttc tcgccgcggg tgcccccaaa gaagcaggcc  | 23540 |
| aacggcacct gcgagcccaa cccccgcctc aacttctacc ccgtctttgc ggtgcccag   | 23600 |
| gcgctggcca cctatcacat cttcttcaag aaccagcga tccccctctc gtgccgcgcc   | 23660 |
| aaccgcaccc gcgccgaccg cctcctgcat ctccgagccg gcgccgccat acctgagatc  | 23720 |
| gcctcccctgg aggaagtccc caagatcttc gaaggtctcg gcaaggacga gaagcgcgcg | 23780 |
| gcaaacgctc tggaaaagaa cgagagcgag ggtcagaacg tgctggtcga gctggaaggc  | 23840 |
| gacaacgcgc gtctggccgt gctcaaacgc accatcgaag tctcccactt cgctacccc   | 23900 |
| gcgctcaacc ttcccccaa ggtcatgcgc tcggctatgg atcagctgct catcaagcgc   | 23960 |
| gccgagcccc tcgagaacga ctccgaggtg gattccgagg acgaaaacc cgtggtctcg   | 24020 |
| gacgaggagc tcgcgcgctg gctgggcacg caggaccccg ccgagttgca agagcggcgc  | 24080 |
| aagatgatga tggggccgt gctggtcacc gccgagctcg agtgccctga gcgcttcttc   | 24140 |
| gccgaccccc agaccctgcg caaggtcgag gagtcctgc actacgcctt ccgccacggc   | 24200 |
| tacgtgcgcc aggcctgcaa gatctccaac gtggagctta gcaacctggt ctctacatg   | 24260 |
| ggcatcctgc acgagaacgg cctcgggcag aacgtcctcc actgcacctt gaccggggag  | 24320 |
| gcccgcgcg actacgtccg cgactgcacg tacctctttc tcaccctcac ctggcagacc   | 24380 |
| gccatggggg tctggcagca gtgtctggag gagcgcaacc ttcgcgagct cgacaagcta  | 24440 |
| ctgagcccg agcggccgca gctctggacg gctttcagcg agcgcaccgc cgcctgccgt   | 24500 |
| ctggccgacc tcatcttccc cgagcgactc aggcaaacc tccagaacgg cctgcccgac   | 24560 |
| ttgtcagcc agagcatgct gcaaaacttt cgctcctca tcctggagcg atccggcatc    | 24620 |
| ttgcccgcca tgagctgcgc cctgcctcc gatttctcc cccttatta tcgcgagtg      | 24680 |
| ccccgcgcg tctggagcca ctgctacctg ctgctctgg ccaactacct cgcccaccac    | 24740 |
| tccgacctca tggaaagactc cagcggcgag gggctgctgg agtgccactg ccgctgcaac | 24800 |
| ctctgcaccc cccaccgctc gctggtctgc aacaccgagc tgctcagcga gacgcaagt   | 24860 |
| atcggtagct ttgagatcca gggaccagag gggccggagg gtgcttccaa cctcaagctc  | 24920 |
| agcccggcgc tctggacttc cgcctacctg cgcaaattta tccccgagga ctatcacgcc  | 24980 |
| caccagatcc aattctacga agaccaatcg cgacccccca aagccccctt cacggcctgt  | 25040 |
| gtcatcacc agagccagat tctggcccaa ttgcaagcca tccagcaggc ccgccaagag   | 25100 |
| ttctcctga aaaaggttca cgggtctat ctggacccc agaccggcga ggaactcaac     | 25160 |
| accccgctac cctccgcgcg cgctctgtgc cgcccgcaga acctgcccgc ccaaggga    | 25220 |
| caagcaggcc atcgcccagc ggcgggccc aaagcagcaa gagctccagg agcagtgga    | 25280 |
| cgaggagtcc tgggacagcc aggcggagga agtctcagac gagggaggag acatggagag  | 25340 |
| ctgggacagc ctagacgag aggaggaggc cgaggagcta gaggacgagc ctctcagga    | 25400 |
| ggaagagccc agcagcgccg cggcaccatc ggcttccaaa gaagcggctc ggagccggcc  | 25460 |
| ggccccgaag cagcagaagc agcaaacgcc gccaccgtcg cccccgacgc caccaccagg  | 25520 |
| ctcactcaaa gccagccgta ggtgggacgc ggtgtccatc gcgggatcgc ccaagcccc   | 25580 |
| agtcgtaag ccaccgggc ggtcgcggcg ggggtactgt tcctggcgcc cccacaagag    | 25640 |



-continued

---

|  |         |
|--|---------|
| caagatcgtc gcctgcctcc agcaactgccg gggcaacatc tccttcgcgc ggcgctactt | 25700   |
| gctcttccac gacgggggtg cggtgccgcg caacgtcctc tactattacc gtcattctcta | 25760   |
| cagcccctac gagacagaag gcccgccctc cgcgtaagac cagccgccag acggtctcct  | 25820   |
| ccgccatcgc gaccgccag gactcggccg ccacgcagga gctcagaaaa cgcattcttc   | 25880   |
| ccaccctgta tgctatcttc cagcagagcc gcggccagca gctggaactg aaagtaaaaa  | 25940   |
| accgtccctc gcgttcgctc acccgagct gtctgtacca caggaggaa gaccaactgc    | 26000   |
| agcgcacgct cgaggacgcc gaggcactgt tcaataaata ctgctcggtg tctcttaagg  | 26060   |
| actgaaagcc cgcgcttttt cagaggctca ttacgtcatc atcatcatga gcaaggacat  | 26120   |
| tcccacgcct tacatgtgga gctaccagcc gcagatggga ctggcggccg gcgcctccca  | 26180   |
| ggattactcc agtcgcatga actggctgag tgccggcccc cacatgatcg ggcgggtcaa  | 26240   |
| tgggattcgt gccaccgcga atcagatact gctggaacag gccgccctca cctccacccc  | 26300   |
| gcgacgtcag ctgaaccgcg ccgcttgccc gcgccccag gtgtaccagg aaaacccgcg   | 26360   |
| cccgaccaca gtctctctgc cagcgcagcgc ggaggccgaa gtcagatga ctaactccgg  | 26420   |
| ggcgcaatta gcggggcggc cccgccacgt cgtcgtccc gggtagagag gtcggcccgc   | 26480   |
| accctacccc tccggcccta taaagaggct gatcattcga ggccgaggta tccagctcaa  | 26540   |
| cgacgaggtg gtgagctcct cgaccggtct tcggcccagc ggagtcttcc agcttgagg   | 26600   |
| cgccggccgc tcttcttca ccaactcgcca ggcctacctg acgctccaga gctcttctc   | 26660   |
| ccagcctcgc tccggcgga tcggcacct ccagttcgtg gaggagttcg tgccctcggg    | 26720   |
| ctacttcaac ccgttctccg gctctcccgg ccgctaccgg gacagcttca tcccacta    | 26780   |
| cgacgcggtg agcgaatccg tggacggcta cgattgatga ccgatggtgc ggcgtaact   | 26840   |
| gcgcgccggc aacatctgca tcaactgccat cgtcctcggg gcttcgcccg ggaggcctgt | 26900   |
| gagttcatct acttccagct cgcgccggac cagcttcagg gcccttcgca cggcgtaag   | 26960   |
| ctcgtgatag aggaagagct cgagagtagc tgcctgcgct gttttacctc gcgcccatc   | 27020   |
| ctagtcgaga gggaacgcgg taggaccacc ctcaccctct actgcactcg tgactccccg  | 27080   |
| gaattacatg aagatctgtg ttgccttcta tgtgccgaac aataaccct cttgtaacta   | 27140   |
| cctacatcca caataacca gaatttgaa actcctttcg tttgttgca g atg aaa      | 27197   |
|  | Met Lys |
| cgc gcc cgc ctc gac gac gac ttc aac ccc gtc tac ccc tat gac        | 27242   |
| Arg Ala Arg Leu Asp Asp Asp Phe Asn Pro Val Tyr Pro Tyr Asp        |         |
| 1430 1435 1440   |         |
| act ccc aac gct ccc tct gtt ccc ttc atc act cct ccc ttc gtc        | 27287   |
| Thr Pro Asn Ala Pro Ser Val Pro Phe Ile Thr Pro Pro Phe Val        |         |
| 1445 1450 1455   |         |
| tcc tcg gac ggc ttg caa gaa aaa cca ccc gga atg ctc agt ctc        | 27332   |
| Ser Ser Asp Gly Leu Gln Glu Lys Pro Pro Gly Met Leu Ser Leu        |         |
| 1460 1465 1470   |         |
| aac tac caa gat cct att acc acc caa aac ggg gca tta act cta        | 27377   |
| Asn Tyr Gln Asp Pro Ile Thr Thr Gln Asn Gly Ala Leu Thr Leu        |         |
| 1475 1480 1485   |         |
| aag ctt ggc agc gga ctg aac ata aac caa gat ggg gaa ctt acc        | 27422   |
| Lys Leu Gly Ser Gly Leu Asn Ile Asn Gln Asp Gly Glu Leu Thr        |         |
| 1490 1495 1500   |         |
| tca gac gcc agc gtt ctc gtc act ccc ccc att aca aaa gcc aac        | 27467   |
| Ser Asp Ala Ser Val Leu Val Thr Pro Pro Ile Thr Lys Ala Asn        |         |
| 1505 1510 1515   |         |

-continued

|             |                     |                     |         |       |
|-------------|---------------------|---------------------|---------|-------|
| aac aca ata | ggc cta gcc ttc aat | gca cct ctt acc ttg | caa agc | 27512 |
| Asn Thr Ile | Gly Leu Ala Phe Asn | Ala Pro Leu Thr Leu | Gln Ser |       |
| 1520        | 1525                | 1530                |         |       |
| gat act tta | aat ctt gct tgt aac | gcc cca ctt acc gtg | caa gac | 27557 |
| Asp Thr Leu | Asn Leu Ala Cys Asn | Ala Pro Leu Thr Val | Gln Asp |       |
| 1535        | 1540                | 1545                |         |       |
| aat agg ttg | gga ata aca tac aac | tct ccc ctc acc ttg | caa aac | 27602 |
| Asn Arg Leu | Gly Ile Thr Tyr Asn | Ser Pro Leu Thr Leu | Gln Asn |       |
| 1550        | 1555                | 1560                |         |       |
| agc gaa ctt | gcc cta gcg gtc acc | ccg cct ctt gac act | gcc aat | 27647 |
| Ser Glu Leu | Ala Leu Ala Val Thr | Pro Pro Leu Asp Thr | Ala Asn |       |
| 1565        | 1570                | 1575                |         |       |
| aac aca ctt | gcg ctt aaa acc gcc | cgg cct ata att aca | aac tct | 27692 |
| Asn Thr Leu | Ala Leu Lys Thr Ala | Arg Pro Ile Ile Thr | Asn Ser |       |
| 1580        | 1585                | 1590                |         |       |
| aat aac gag | ctt aca ctc tcc gct | gat gct ccc cta aac | acc agc | 27737 |
| Asn Asn Glu | Leu Thr Leu Ser Ala | Asp Ala Pro Leu Asn | Thr Ser |       |
| 1595        | 1600                | 1605                |         |       |
| acg ggt acc | ctc cgc cta caa agc | gca gca cca ctg ggg | cta gtt | 27782 |
| Thr Gly Thr | Leu Arg Leu Gln Ser | Ala Ala Pro Leu Gly | Leu Val |       |
| 1610        | 1615                | 1620                |         |       |
| gac caa acc | ctg cga gtg ctt ttt | tct aac cca ctc tac | ttg caa | 27827 |
| Asp Gln Thr | Leu Arg Val Leu Phe | Ser Asn Pro Leu Tyr | Leu Gln |       |
| 1625        | 1630                | 1635                |         |       |
| aac aac ttt | ctc tca cta gcc att | gaa cgc cca ttg gct | tta act | 27872 |
| Asn Asn Phe | Leu Ser Leu Ala Ile | Glu Arg Pro Leu Ala | Leu Thr |       |
| 1640        | 1645                | 1650                |         |       |
| acc act ggt | tct atg gct atg cag | att tcc caa cca tta | aaa gtg | 27917 |
| Thr Thr Gly | Ser Met Ala Met Gln | Ile Ser Gln Pro Leu | Lys Val |       |
| 1655        | 1660                | 1665                |         |       |
| gaa gac gga | agc tta agc ttg agc | att gaa agc cct cta | aat cta | 27962 |
| Glu Asp Gly | Ser Leu Ser Leu Ser | Ile Glu Ser Pro Leu | Asn Leu |       |
| 1670        | 1675                | 1680                |         |       |
| aaa aac gga | aat ctt act tta gga | acc caa agt ccc cta | act gtc | 28007 |
| Lys Asn Gly | Asn Leu Thr Leu Gly | Thr Gln Ser Pro Leu | Thr Val |       |
| 1685        | 1690                | 1695                |         |       |
| act ggt aac | aac ctc agc ctt aca | aca aca gcc cca tta | acg gtt | 28052 |
| Thr Gly Asn | Asn Leu Ser Leu Thr | Thr Thr Ala Pro Leu | Thr Val |       |
| 1700        | 1705                | 1710                |         |       |
| cag aac aac | gct cta gcc ctc tca | gtg tta ctg ccg ctt | aga cta | 28097 |
| Gln Asn Asn | Ala Leu Ala Leu Ser | Val Leu Leu Pro Leu | Arg Leu |       |
| 1715        | 1720                | 1725                |         |       |
| ttt aat aac | acc tca ctg gga gtg | gca ttc aac cca ccc | att tct | 28142 |
| Phe Asn Asn | Thr Ser Leu Gly Val | Ala Phe Asn Pro Pro | Ile Ser |       |
| 1730        | 1735                | 1740                |         |       |
| tca gca aac | aac ggg ctg tct ctt | gac att gga aat ggc | ctt aca | 28187 |
| Ser Ala Asn | Asn Gly Leu Ser Leu | Asp Ile Gly Asn Gly | Leu Thr |       |
| 1745        | 1750                | 1755                |         |       |
| ctg caa tac | aac agg ctc gta gtg | aac att ggc ggc ggg | cta cag | 28232 |
| Leu Gln Tyr | Asn Arg Leu Val Val | Asn Ile Gly Gly Gly | Leu Gln |       |
| 1760        | 1765                | 1770                |         |       |
| ttt aac aac | ggt gct att acc gct | tcc ata aat gca gct | ctg ccg | 28277 |
| Phe Asn Asn | Gly Ala Ile Thr Ala | Ser Ile Asn Ala Ala | Leu Pro |       |
| 1775        | 1780                | 1785                |         |       |
| ttg cag tat | tcc aat aac cag ctt | tct ctt aat att gga | ggc ggg | 28322 |
| Leu Gln Tyr | Ser Asn Asn Gln Leu | Ser Leu Asn Ile Gly | Gly Gly |       |
| 1790        | 1795                | 1800                |         |       |

-continued

|  |       |
|--|-------|
| ctg cga tac aac ggc act tac aaa aat tta gcc gtc aaa acc gac        | 28367 |
| Leu Arg Tyr Asn Gly Thr Tyr Lys Asn Leu Ala Val Lys Thr Asp        |       |
| 1805 1810 1815   |       |
| tct ttt agg ggt ctt gaa att gac agt aat cag ttc ctg gtg cca        | 28412 |
| Ser Phe Arg Gly Leu Glu Ile Asp Ser Asn Gln Phe Leu Val Pro        |       |
| 1820 1825 1830   |       |
| aga ctg ggt tct ggt cta aag ttt gat caa tat ggg tac att agc        | 28457 |
| Arg Leu Gly Ser Gly Leu Lys Phe Asp Gln Tyr Gly Tyr Ile Ser        |       |
| 1835 1840 1845   |       |
| gtc ata cct cca act gtt acg cca aca aca ctt tgg act aca gca        | 28502 |
| Val Ile Pro Pro Thr Val Thr Pro Thr Thr Leu Trp Thr Thr Ala        |       |
| 1850 1855 1860   |       |
| gac cct tct ccc aac gct act ttt tac gac agc tta gat gct aag        | 28547 |
| Asp Pro Ser Pro Asn Ala Thr Phe Tyr Asp Ser Leu Asp Ala Lys        |       |
| 1865 1870 1875   |       |
| gta tgg ctg gcc tta gta aaa tgc aac ggc atg gtt aat gga acc        | 28592 |
| Val Trp Leu Ala Leu Val Lys Cys Asn Gly Met Val Asn Gly Thr        |       |
| 1880 1885 1890   |       |
| ata gcc ata aag gct tta aaa ggt act ctg ctc caa cct acg gct        | 28637 |
| Ile Ala Ile Lys Ala Leu Lys Gly Thr Leu Leu Gln Pro Thr Ala        |       |
| 1895 1900 1905   |       |
| agt ttt att tct ttt gtt atg tat ttt tac agc aat ggc acc aga        | 28682 |
| Ser Phe Ile Ser Phe Val Met Tyr Phe Tyr Ser Asn Gly Thr Arg        |       |
| 1910 1915 1920   |       |
| aga act aac tac ccc acg ttt gaa aat gaa ggc ata cta gct agt        | 28727 |
| Arg Thr Asn Tyr Pro Thr Phe Glu Asn Glu Gly Ile Leu Ala Ser        |       |
| 1925 1930 1935   |       |
| agt gct aca tgg ggt tat cgt caa gga aac tcg gca aac acc aac        | 28772 |
| Ser Ala Thr Trp Gly Tyr Arg Gln Gly Asn Ser Ala Asn Thr Asn        |       |
| 1940 1945 1950   |       |
| gtc acc agt gcc gtt gaa ttt atg cct agc tcc aca aga tat cct        | 28817 |
| Val Thr Ser Ala Val Glu Phe Met Pro Ser Ser Thr Arg Tyr Pro        |       |
| 1955 1960 1965   |       |
| gtt aac aag ggt act gag gtt cag aac atg gaa ctc acc tac act        | 28862 |
| Val Asn Lys Gly Thr Glu Val Gln Asn Met Glu Leu Thr Tyr Thr        |       |
| 1970 1975 1980   |       |
| ttc ttg cag gga gac ccc act atg gcc ata tca ttt caa gct att        | 28907 |
| Phe Leu Gln Gly Asp Pro Thr Met Ala Ile Ser Phe Gln Ala Ile        |       |
| 1985 1990 1995   |       |
| tat aac cat gct ttg gaa ggt tac tct tta aaa ttt acc tgg cga        | 28952 |
| Tyr Asn His Ala Leu Glu Gly Tyr Ser Leu Lys Phe Thr Trp Arg        |       |
| 2000 2005 2010   |       |
| gtt cgc aac agg gaa cgc ttt gat atc ccc tgc tgt tct ttt tct        | 28997 |
| Val Arg Asn Arg Glu Arg Phe Asp Ile Pro Cys Cys Ser Phe Ser        |       |
| 2015 2020 2025   |       |
| tac ata acg gaa gaa taa acactgtttt tcttttcaat gtttttattc           | 29045 |
| Tyr Ile Thr Glu Glu  |       |
| 2030   |       |
| tgctttttta cacagttcga accgtcagac tccctcccc cttccacttc acccgtaca    | 29105 |
| cctcccgcct cccctggatc gctgcgtaca actgcagttt ggtgttcaga cacgggttct  | 29165 |
| taggtgacag tatccacacg gcctctttgc cggccaggcg ctggtccgta atgctcacia  | 29225 |
| atccctccga cacgtctctc agacacacgg tggaatccaa ggcgccgctc tacaaaacia  | 29285 |
| acacagtcac gctctccacg gttctctctc tcggtctgtac tgcgccagcg tgaacgggcg | 29345 |
| atggtgctcc atcagggtctc gcagcaaccg ctgtcggcgc ggctcaccca ggctccggcg | 29405 |

---

-continued

---

aaaagcgccc cgtctgggag tgctattcaa aaaacgcacc gcctttatca acagtctcct 29465  
cgtgcgggcg gcgcagcagc gcacctggat ctctgtcagg tctttacaat aggtacagcc 29525  
catcaccacc atgttgttta aaatcccaaa gctaaacacg ctccacccaa atgacatgaa 29585  
ttccagcacc gccgcggcgt ggccatcata caatatgcgg aggtaaatca ggtgccgccc 29645  
cctaatacaa acgctcccca tatacatcac ctcccttaggc agttgataat taaccacctc 29705  
ccggtaccag gaaacctca cgtttactaa agccccaac accaacattt taaaccagtt 29765  
agccagcacc accctcccgc cttacactg cagcgacccc ggctgtttac aatgacagt 29825  
aatcaccacc ctctcatacc ccctaatgac ctggcgtggc tccacatcta tagtagcaca 29885  
gcacacgcac accctcatgt aatgcttcat cacaatctt tccaagggg ttagtatcat 29945  
gtcccagggt acgggccact cctgcagcac ggtgaaagg acgcaggcgg gaacagtcct 30005  
cacctcggac acataatgca tattcagatg ttcacactct aaaaccccgg ggcttccctc 30065  
caacgcagcc actggcaagt tctcagaggg tgggtgaagg cgggtggtgct gatagggact 30125  
caatctgtgt cgacaccgtc tgcgcgttg catcgtagac caacgcttg cgcaccgcct 30185  
cgtacttcgc ccaaagaaaa cgggtgcgac gccaacacac ttcgcgtac cgtgggttcc 30245  
gcactcgagc tcgctcagtt ctcaacgcac aatgcagcca ttcctgtaac ccacacaaca 30305  
gtcgtcggc ttccaaagag atgtgcacct cgtatcttat aacgtcccga tatatatcca 30365  
agcaggcagt cagggccact tgcaaccagt gcacgcaggc ggactgatcg cgacacactg 30425  
gaggtggagg gagagacgga agaggcatgt tactccagac ggtcgaaaag cggatcaaag 30485  
tgcagatcgc gaagatggca gcgatcccgc ccgctacgct ggtgatagat cacagccagg 30545  
tcaaacataa tgcggttttc caaatgacct attaccgct ccaccagagc cgccacgcgc 30605  
acttccagaa acaccagcac ggctacggca ttctcctcaa aatcttcaa cattaagctg 30665  
catgattgaa tcaccccaa ataattctcc tccttccatt ctgcgaaaat ttgagtaaaa 30725  
acctctcgca gattagctcc gtggcgttca aaaaggtcac ttagagcgcc ctccaccgcc 30785  
atgcgcaagc acaccctcat gattgaaaaa tgccagtctc ctgaaccacc tgcagttgat 30845  
ttaaagacc tatattagga tcaattccac tctcccagc ctccacgctg agcattagct 30905  
gcaaaaagtc atttaaatct tcgcaaaacta gcgcgtaag ctccgcccgc ggaattaggt 30965  
ctgaagcagt caccacacac ataatttcca gtgaaggagt cagtctaagc agcaaaaagc 31025  
cgcatgagca gtgttgaaaa ggagggtca cgcaatgtaa catatgcagc caaaaatctc 31085  
caaggtgtct gtgcataaac tccaccactg aaaagtccaa atcatgtaa tatgccatca 31145  
ccgcctcagg aaccaccagc gacacaaaaa cgggccgtag caaatacatg gtgtcctgca 31205  
aagcaaaaac acatttatac catagaggcg cgaattactt ggggaaaaat cactcgtctc 31265  
aaaactaac aggccaccgt ctgacocgc cagccataaa aaaagcggtt cgaatgatta 31325  
aaaagaataa tagacaccto ccaccagta ctccgctgca actcgtgcgc ccctatcaaa 31385  
accccgcgga cgttcatgtc ggccatagaa aaaatgcggc ccaaatatcc caccggaatc 31445  
tccacggcca gctgcagtga tagcaaaaga acgcatgag gagcaatcac aaaatttca 31505  
ggcgataaaa gcacataaag gttagaatag ccctgctgca caggtaataa agcccgcgag 31565  
ctcagcaaat gcacataaac cgcttcagcc atcccgtett accgcgaaca aaaggctcac 31625  
agtacacagt tactcaaccc acacgccaca cagtatttat acactcctca atcggcacgt 31685

-continued

---

```

caccgcgccc gaacaaactc caaaagtcca aaaagtccaa aacgcccgcg taaaagcccc 31745
caaaaacagc acttcctcat ttactctccc acagtacgtc acttccgccc cgcccgcgcg 31805
cctcgccccg cctcaccctc cgcgctccac cccgcgcccc acgtcagact cccaccccgc 31865
cccgcgcccc cgtcatccgc accccacctc cactccaccc ctaaccccgc ctctcatta 31925
tcataattggc accgttccca aataaggtat attatgatga tg 31967

```

```

<210> SEQ ID NO 13
<211> LENGTH: 508
<212> TYPE: PRT
<213> ORGANISM: Simian adenovirus

```

```

<400> SEQUENCE: 13

```

```

Met Arg Arg Ala Val Gly Val Pro Pro Val Met Ala Tyr Ala Glu Gly
1          5          10          15
Pro Pro Pro Ser Tyr Glu Thr Val Met Gly Ala Ala Asp Ser Pro Ala
20          25          30
Thr Leu Glu Ala Leu Tyr Val Pro Pro Arg Tyr Leu Gly Pro Thr Glu
35          40          45
Gly Arg Asn Ser Ile Arg Tyr Ser Glu Leu Ala Pro Leu Tyr Asp Thr
50          55          60
Thr Arg Val Tyr Leu Val Asp Asn Lys Ser Ala Asp Ile Ala Ser Leu
65          70          75
Asn Tyr Gln Asn Asp His Ser Asn Phe Leu Thr Thr Val Val Gln Asn
85          90          95
Asn Asp Phe Thr Pro Val Glu Ala Gly Thr Gln Thr Ile Asn Phe Asp
100         105         110
Glu Arg Ser Arg Trp Gly Gly Asp Leu Lys Thr Ile Leu Arg Thr Asn
115         120         125
Met Pro Asn Ile Asn Glu Phe Met Ser Thr Asn Lys Phe Arg Ala Arg
130         135         140
Leu Met Val Glu Lys Val Asn Lys Glu Thr Asn Ala Pro Arg Tyr Glu
145         150         155
Trp Phe Glu Phe Thr Leu Pro Glu Gly Asn Tyr Ser Glu Thr Met Thr
165         170         175
Ile Asp Leu Met Asn Asn Ala Ile Val Asp Asn Tyr Leu Glu Val Gly
180         185         190
Arg Gln Asn Gly Val Leu Glu Ser Asp Ile Gly Val Lys Phe Asp Thr
195         200         205
Arg Asn Phe Arg Leu Gly Trp Asp Pro Val Thr Lys Leu Val Met Pro
210         215         220
Gly Val Tyr Thr Asn Glu Ala Phe His Pro Asp Ile Val Leu Leu Pro
225         230         235
Gly Cys Gly Val Asp Phe Thr Gln Ser Arg Leu Ser Asn Leu Leu Gly
245         250         255
Ile Arg Lys Arg Met Pro Phe Gln Ala Gly Phe Gln Ile Met Tyr Glu
260         265         270
Asp Leu Glu Gly Gly Asn Ile Pro Ala Leu Leu Asp Val Ala Lys Tyr
275         280         285
Glu Ala Ser Ile Gln Lys Ala Arg Glu Gln Gly Gln Glu Ile Arg Gly
290         295         300

```

-continued

---

Asp Asn Phe Thr Val Ile Pro Arg Asp Val Glu Ile Val Pro Val Glu  
 305 310 315 320  
 Lys Asp Ser Lys Asp Arg Ser Tyr Asn Leu Leu Pro Gly Asp Gln Thr  
 325 330 335  
 Asn Thr Ala Tyr Arg Ser Trp Phe Leu Ala Tyr Asn Tyr Gly Asp Pro  
 340 345 350  
 Glu Lys Gly Val Arg Ser Trp Thr Leu Leu Thr Thr Thr Asp Val Thr  
 355 360 365  
 Cys Gly Ser Gln Gln Val Tyr Trp Ser Leu Pro Asp Met Met Gln Asp  
 370 375 380  
 Pro Val Thr Phe Arg Pro Ser Ser Gln Val Ser Asn Tyr Pro Val Val  
 385 390 395 400  
 Gly Val Glu Leu Leu Pro Val His Ala Lys Ser Phe Tyr Asn Glu Gln  
 405 410 415  
 Ala Val Tyr Ser Gln Leu Ile Arg Gln Ser Thr Ala Leu Thr His Val  
 420 425 430  
 Phe Asn Arg Phe Pro Glu Asn Gln Ile Leu Val Arg Pro Pro Ala Pro  
 435 440 445  
 Thr Ile Thr Thr Val Ser Glu Asn Val Pro Ala Leu Thr Asp His Gly  
 450 455 460  
 Thr Leu Pro Leu Arg Ser Ser Ile Ser Gly Val Gln Arg Val Thr Ile  
 465 470 475 480  
 Thr Asp Ala Arg Arg Arg Thr Cys Pro Tyr Val His Lys Ala Leu Gly  
 485 490 495  
 Ile Val Ala Pro Lys Val Leu Ser Ser Arg Thr Phe  
 500 505

&lt;210&gt; SEQ ID NO 14

&lt;211&gt; LENGTH: 917

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Simian adenovirus

&lt;400&gt; SEQUENCE: 14

Met Ala Thr Pro Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ala  
 1 5 10 15  
 Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala  
 20 25 30  
 Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro  
 35 40 45  
 Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu  
 50 55 60  
 Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr  
 65 70 75 80  
 Lys Val Arg Phe Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met  
 85 90 95  
 Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser  
 100 105 110  
 Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly  
 115 120 125  
 Ala Pro Asn Pro Ser Glu Trp Lys Gly Ser Asp Asn Lys Ile Ser Val  
 130 135 140  
 Arg Gly Gln Ala Pro Phe Phe Ser Thr Ser Ile Thr Lys Asp Gly Ile  
 145 150 155 160

-continued

---

Gln Val Ala Thr Asp Thr Ser Ser Gly Ala Val Tyr Ala Lys Lys Glu  
 165 170 175  
 Tyr Gln Pro Glu Pro Gln Val Gly Gln Glu Gln Trp Asn Ser Glu Ala  
 180 185 190  
 Ser Asp Ser Asp Lys Val Ala Gly Arg Ile Leu Lys Asp Thr Thr Pro  
 195 200 205  
 Met Phe Pro Cys Tyr Gly Ser Tyr Ala Lys Pro Thr Asn Glu Gln Gly  
 210 215 220  
 Gly Gln Gly Thr Asn Thr Val Asp Leu Gln Phe Phe Ala Ser Ser Ser  
 225 230 235 240  
 Ala Thr Ser Thr Pro Lys Ala Val Leu Tyr Ala Glu Asp Val Ala Ile  
 245 250 255  
 Glu Ala Pro Asp Thr His Leu Val Tyr Lys Pro Ala Val Thr Thr Thr  
 260 265 270  
 Thr Thr Ser Ser Gln Asp Leu Leu Thr Gln Gln Ala Ala Pro Asn Arg  
 275 280 285  
 Pro Asn Tyr Ile Gly Phe Arg Asp Asn Phe Ile Gly Leu Met Tyr Tyr  
 290 295 300  
 Asn Ser Thr Gly Asn Met Gly Val Leu Ala Gly Gln Ala Ser Gln Leu  
 305 310 315 320  
 Asn Ala Val Val Asp Leu Gln Asp Arg Asn Thr Glu Leu Ser Tyr Gln  
 325 330 335  
 Leu Met Leu Asp Ala Leu Gly Asp Arg Ser Arg Tyr Phe Ser Met Trp  
 340 345 350  
 Asn Gln Ala Val Asp Ser Tyr Asp Pro Asp Val Arg Ile Ile Glu Asn  
 355 360 365  
 His Gly Val Glu Asp Glu Leu Pro Asn Tyr Cys Phe Pro Leu Gly Gly  
 370 375 380  
 Ser Leu Val Thr Glu Thr Tyr Thr Gly Leu Ser Pro Gln Asn Gly Ser  
 385 390 395 400  
 Asn Thr Trp Thr Thr Asp Ser Thr Thr Tyr Ala Thr Arg Gly Val Glu  
 405 410 415  
 Ile Gly Ser Gly Asn Met Phe Ala Met Glu Ile Asn Leu Ala Ala Asn  
 420 425 430  
 Leu Trp Arg Ser Phe Leu Tyr Ser Asn Val Ala Leu Tyr Leu Pro Asp  
 435 440 445  
 Glu Tyr Lys Leu Thr Pro Asp Asn Ile Thr Leu Pro Asp Asn Lys Asn  
 450 455 460  
 Thr Tyr Asp Tyr Met Asn Gly Arg Val Ala Ala Pro Ser Ser Leu Asp  
 465 470 475 480  
 Thr Tyr Val Asn Ile Gly Ala Arg Trp Ser Pro Asp Pro Met Asp Asn  
 485 490 495  
 Val Asn Pro Phe Asn His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser  
 500 505 510  
 Met Leu Leu Gly Asn Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro  
 515 520 525  
 Gln Lys Phe Phe Ala Ile Lys Asn Leu Leu Leu Leu Pro Gly Ser Tyr  
 530 535 540  
 Thr Tyr Glu Trp Asn Phe Arg Lys Asp Val Asn Met Ile Leu Gln Ser  
 545 550 555 560

-continued

```

Ser Leu Gly Asn Asp Leu Arg Val Asp Gly Ala Ser Val Arg Phe Asp
           565                               570                   575

Ser Ile Asn Leu Tyr Ala Asn Phe Phe Pro Met Ala His Asn Thr Ala
           580                               585                   590

Ser Thr Leu Glu Ala Met Leu Arg Asn Asp Thr Asn Asp Gln Ser Phe
           595                               600                   605

Asn Asp Tyr Leu Cys Ala Ala Asn Met Leu Tyr Pro Ile Pro Ala Asn
           610                               615                   620

Ala Thr Ser Val Pro Ile Ser Ile Pro Ser Arg Asn Trp Ala Ala Phe
           625                               630                   635                   640

Arg Gly Trp Ser Phe Thr Arg Leu Lys Thr Lys Glu Thr Pro Ser Leu
           645                               650                   655

Gly Ser Gly Phe Asp Pro Tyr Phe Thr Tyr Ser Gly Ser Ile Pro Tyr
           660                               665                   670

Leu Asp Gly Thr Phe Tyr Leu Asn His Thr Phe Lys Lys Val Ser Ile
           675                               680                   685

Met Phe Asp Ser Ser Val Ser Trp Pro Gly Asn Asp Arg Leu Leu Thr
           690                               695                   700

Pro Asn Glu Phe Glu Ile Lys Arg Thr Val Asp Gly Glu Gly Tyr Asn
           705                               710                   715                   720

Val Ala Gln Cys Asn Met Thr Lys Asp Trp Phe Leu Ile Gln Met Leu
           725                               730                   735

Ser His Tyr Asn Ile Gly Tyr Gln Gly Phe Tyr Val Pro Glu Gly Tyr
           740                               745                   750

Lys Asp Arg Met Tyr Ser Phe Phe Arg Asn Phe Gln Pro Met Ser Arg
           755                               760                   765

Gln Val Val Asp Thr Thr Thr Tyr Thr Asp Tyr Lys Asn Val Thr Leu
           770                               775                   780

Pro Phe Gln His Asn Asn Ser Gly Phe Val Gly Tyr Met Gly Pro Thr
           785                               790                   795                   800

Met Arg Glu Gly Gln Ala Tyr Pro Ala Asn Tyr Pro Tyr Pro Leu Ile
           805                               810                   815

Gly Lys Thr Ala Val Pro Ser Leu Thr Gln Lys Lys Phe Leu Cys Asp
           820                               825                   830

Arg Thr Met Trp Arg Ile Pro Phe Ser Ser Asn Phe Met Ser Met Gly
           835                               840                   845

Ala Leu Thr Asp Leu Gly Gln Asn Met Leu Tyr Ala Asn Ser Ala His
           850                               855                   860

Ala Leu Asp Met Thr Phe Glu Val Asp Pro Met Asp Glu Pro Thr Leu
           865                               870                   875                   880

Leu Tyr Val Leu Phe Glu Val Phe Asp Val Val Arg Ile His Gln Pro
           885                               890                   895

His Arg Gly Val Ile Glu Ala Val Tyr Leu Arg Thr Pro Phe Ser Ala
           900                               905                   910

Gly Asn Ala Thr Thr
           915

```

&lt;210&gt; SEQ ID NO 15

&lt;211&gt; LENGTH: 607

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Simian adenovirus

&lt;400&gt; SEQUENCE: 15



-continued

---

Met Lys Arg Ala Arg Leu Asp Asp Asp Phe Asn Pro Val Tyr Pro Tyr  
 1 5 10 15  
 Asp Thr Pro Asn Ala Pro Ser Val Pro Phe Ile Thr Pro Pro Phe Val  
 20 25 30  
 Ser Ser Asp Gly Leu Gln Glu Lys Pro Pro Gly Met Leu Ser Leu Asn  
 35 40 45  
 Tyr Gln Asp Pro Ile Thr Thr Gln Asn Gly Ala Leu Thr Leu Lys Leu  
 50 55 60  
 Gly Ser Gly Leu Asn Ile Asn Gln Asp Gly Glu Leu Thr Ser Asp Ala  
 65 70 75 80  
 Ser Val Leu Val Thr Pro Pro Ile Thr Lys Ala Asn Asn Thr Ile Gly  
 85 90 95  
 Leu Ala Phe Asn Ala Pro Leu Thr Leu Gln Ser Asp Thr Leu Asn Leu  
 100 105 110  
 Ala Cys Asn Ala Pro Leu Thr Val Gln Asp Asn Arg Leu Gly Ile Thr  
 115 120 125  
 Tyr Asn Ser Pro Leu Thr Leu Gln Asn Ser Glu Leu Ala Leu Ala Val  
 130 135 140  
 Thr Pro Pro Leu Asp Thr Ala Asn Asn Thr Leu Ala Leu Lys Thr Ala  
 145 150 155 160  
 Arg Pro Ile Ile Thr Asn Ser Asn Asn Glu Leu Thr Leu Ser Ala Asp  
 165 170 175  
 Ala Pro Leu Asn Thr Ser Thr Gly Thr Leu Arg Leu Gln Ser Ala Ala  
 180 185 190  
 Pro Leu Gly Leu Val Asp Gln Thr Leu Arg Val Leu Phe Ser Asn Pro  
 195 200 205  
 Leu Tyr Leu Gln Asn Asn Phe Leu Ser Leu Ala Ile Glu Arg Pro Leu  
 210 215 220  
 Ala Leu Thr Thr Thr Gly Ser Met Ala Met Gln Ile Ser Gln Pro Leu  
 225 230 235 240  
 Lys Val Glu Asp Gly Ser Leu Ser Leu Ser Ile Glu Ser Pro Leu Asn  
 245 250 255  
 Leu Lys Asn Gly Asn Leu Thr Leu Gly Thr Gln Ser Pro Leu Thr Val  
 260 265 270  
 Thr Gly Asn Asn Leu Ser Leu Thr Thr Thr Ala Pro Leu Thr Val Gln  
 275 280 285  
 Asn Asn Ala Leu Ala Leu Ser Val Leu Leu Pro Leu Arg Leu Phe Asn  
 290 295 300  
 Asn Thr Ser Leu Gly Val Ala Phe Asn Pro Pro Ile Ser Ser Ala Asn  
 305 310 315 320  
 Asn Gly Leu Ser Leu Asp Ile Gly Asn Gly Leu Thr Leu Gln Tyr Asn  
 325 330 335  
 Arg Leu Val Val Asn Ile Gly Gly Gly Leu Gln Phe Asn Asn Gly Ala  
 340 345 350  
 Ile Thr Ala Ser Ile Asn Ala Ala Leu Pro Leu Gln Tyr Ser Asn Asn  
 355 360 365  
 Gln Leu Ser Leu Asn Ile Gly Gly Gly Leu Arg Tyr Asn Gly Thr Tyr  
 370 375 380  
 Lys Asn Leu Ala Val Lys Thr Asp Ser Phe Arg Gly Leu Glu Ile Asp  
 385 390 395 400

-continued

---

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ser | Asn | Gln | Phe | Leu | Val | Pro | Arg | Leu | Gly | Ser | Gly | Leu | Lys | Phe | Asp |
|     |     |     |     | 405 |     |     |     |     | 410 |     |     |     |     | 415 |     |
| Gln | Tyr | Gly | Tyr | Ile | Ser | Val | Ile | Pro | Pro | Thr | Val | Thr | Pro | Thr | Thr |
|     |     |     | 420 |     |     |     |     | 425 |     |     |     |     | 430 |     |     |
| Leu | Trp | Thr | Thr | Ala | Asp | Pro | Ser | Pro | Asn | Ala | Thr | Phe | Tyr | Asp | Ser |
|     |     | 435 |     |     |     |     | 440 |     |     |     |     | 445 |     |     |     |
| Leu | Asp | Ala | Lys | Val | Trp | Leu | Ala | Leu | Val | Lys | Cys | Asn | Gly | Met | Val |
|     | 450 |     |     |     |     | 455 |     |     |     |     | 460 |     |     |     |     |
| Asn | Gly | Thr | Ile | Ala | Ile | Lys | Ala | Leu | Lys | Gly | Thr | Leu | Leu | Gln | Pro |
| 465 |     |     |     |     | 470 |     |     |     |     | 475 |     |     |     |     | 480 |
| Thr | Ala | Ser | Phe | Ile | Ser | Phe | Val | Met | Tyr | Phe | Tyr | Ser | Asn | Gly | Thr |
|     |     |     | 485 |     |     |     |     |     | 490 |     |     |     |     | 495 |     |
| Arg | Arg | Thr | Asn | Tyr | Pro | Thr | Phe | Glu | Asn | Glu | Gly | Ile | Leu | Ala | Ser |
|     |     |     | 500 |     |     |     |     | 505 |     |     |     |     |     | 510 |     |
| Ser | Ala | Thr | Trp | Gly | Tyr | Arg | Gln | Gly | Asn | Ser | Ala | Asn | Thr | Asn | Val |
|     |     | 515 |     |     |     |     | 520 |     |     |     |     | 525 |     |     |     |
| Thr | Ser | Ala | Val | Glu | Phe | Met | Pro | Ser | Ser | Thr | Arg | Tyr | Pro | Val | Asn |
|     | 530 |     |     |     |     | 535 |     |     |     |     | 540 |     |     |     |     |
| Lys | Gly | Thr | Glu | Val | Gln | Asn | Met | Glu | Leu | Thr | Tyr | Thr | Phe | Leu | Gln |
| 545 |     |     |     |     | 550 |     |     |     |     | 555 |     |     |     |     | 560 |
| Gly | Asp | Pro | Thr | Met | Ala | Ile | Ser | Phe | Gln | Ala | Ile | Tyr | Asn | His | Ala |
|     |     |     |     | 565 |     |     |     |     | 570 |     |     |     |     | 575 |     |
| Leu | Glu | Gly | Tyr | Ser | Leu | Lys | Phe | Thr | Trp | Arg | Val | Arg | Asn | Arg | Glu |
|     |     |     | 580 |     |     |     |     | 585 |     |     |     |     | 590 |     |     |
| Arg | Phe | Asp | Ile | Pro | Cys | Cys | Ser | Phe | Ser | Tyr | Ile | Thr | Glu | Glu |     |
|     |     | 595 |     |     |     |     | 600 |     |     |     |     | 605 |     |     |     |

---

1. A method of producing a chimeric adenovirus in a selected host cell type, said chimeric adenovirus derived from a parental adenovirus which is not useful for adenovirus virion production in the host cell type, said method comprising the steps of:

- (a) generating a chimeric adenovirus comprising:
- (i) adenovirus sequences of the left terminal end and right terminal end of a first adenovirus which grows in a the selected host cell type, said left end region comprising the 5' inverted terminal repeat (ITRs), and said right end region comprising the 3' inverted terminal repeat (ITRs); and
  - (ii) the internal regions from the parental adenovirus, said internal regions comprising the late genes encoding the penton, hexon, and fiber and lacking its native 5' and 3' terminal regions from the parental adenovirus;

wherein the resulting chimeric adenovirus comprises, from 5' to 3', a left terminal region of the first adenovirus, the internal region of the parental adenovirus, and the right terminal region of the first adenovirus; and

- b) producing said chimeric adenovirus in the presence of functional adenovirus E1a, E1b, and E4 ORF6 genes from the first adenovirus or from an adenovirus serotype which transcomplements the first adenovirus, and

further in the presence of necessary adenoviral structural genes from the left end of the first adenovirus or from an adenovirus serotype which transcomplements the first adenovirus.

2. The method according to claim 1, wherein the internal region of the parental adenovirus further comprises one or more functional adenovirus genes selected from the group consisting of Endoprotease open reading frame, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, protein VIII, pTP (pre-Terminal Protein), 52/55 kDa protein, protein VII, Mu and protein VI.

3. The method according to claim 1, wherein the polymerase, terminal protein and 52/55 kDa protein functions are provided in trans.

4. The method according to claim 1, wherein the first adenovirus further comprises the polymerase, terminal protein and 52/55 kDa protein functions.

5. The method according to claim 1, wherein the chimeric adenovirus comprises the adenoviral late genes 1, 2, 3, 4, and 5 of the parental adenovirus.

6. The method according to claim 1, wherein the selected host cell stably contains one or more functional adenovirus E1a, E1b or E4 ORF6 genes.

7. The method according to claim 1, wherein the chimeric adenovirus comprises one or more functional adenovirus E1a, E1b or E4 ORF6 genes from the first adenovirus.

8. The method according to claim 1, wherein the first adenovirus is of human origin.

9. The method according to claim 1, wherein the first adenovirus is of simian origin.

10. The method according to claim 1, further comprising the step of isolating the chimeric adenovirus.

11. A method for generating a chimeric adenovirus for production in a selected host cell, said chimeric adenovirus derived from a parental adenovirus which is not useful for adenovirus virion production in the host cell type, said method comprising the step of generating a chimeric adenovirus comprising:

5' and 3' terminal regions of a first adenovirus which grows in a selected host cell type, said 5' terminal regions comprising the 5' inverted terminal repeat (ITRs) and necessary E1 gene functions, and said 3' terminal regions comprising inverted terminal repeat (ITRs) and necessary E4 gene functions; and

internal regions from the parental adenovirus, said internal regions comprising the hexon, penton base and fiber and lacking at least its native 5' and 3' terminal regions;

wherein the resulting chimeric adenovirus comprises, from 5' to 3', the 5' terminal region of the first adenovirus, the internal region of the parental adenovirus, and the 3' terminal regions of the first adenovirus.

12. A chimeric adenovirus produced according to the method of claim 1.

13. A chimeric adenovirus comprising a hexon protein of a selected adenovirus serotype, said chimeric adenovirus derived from a parental adenovirus which is not useful for adenovirus virion production in the host cell type, said chimeric adenovirus comprising:

(a) adenovirus sequences of the left terminal end of a first adenovirus which grows in a selected host cell type, said left end region comprising the E1a, E1b and 5' inverted terminal repeat (ITRs);

(b) adenovirus sequences of the internal region of the selected parental adenovirus, said internal region comprising the genes encoding the penton, hexon and fiber of the selected adenovirus;

(c) adenovirus sequences of the right terminal end of the first adenovirus, said right end region comprising the necessary E4 gene functions and the 3' inverted terminal repeat (ITRs),

wherein the resulting chimeric adenovirus comprises adenoviral structural and regulatory proteins necessary for infection and replication.

14. The chimeric adenovirus according to claim 13, wherein the chimeric adenovirus further comprises the IIIa, 52/55kDa and pTP of the selected adenovirus serotype.

15. The chimeric adenovirus according to claim 13, wherein chimeric adenovirus comprises the polymerase of the first adenovirus.

16. The chimeric adenovirus according to claim 13, wherein the chimeric adenovirus expresses a functional chimeric protein formed from the first adenovirus and the selected adenovirus, said chimeric protein is selected from the group consisting of polymerase, terminal protein, 52/55 kDa protein, and IIIa.

17. The chimeric adenovirus according to claim 13, wherein the chimeric adenovirus comprises the terminal protein, 52/55 kDa, and/or IIIa of the selected adenovirus.

18. A host cell in vitro comprising a chimeric adenovirus according to claim 12.

19. The host cell according to claim 18, wherein said host cell is a human cell.

20-43. (canceled)

44. The chimeric adenovirus according to claim 12, wherein said chimeric adenovirus is functionally deleted in E1a and/or E1b.

45. The chimeric adenovirus according to claim 12, wherein said chimeric adenovirus is functionally deleted in E4ORF6.

46. The method according to claim 1, wherein the parental adenovirus is selected from the group consisting of adenovirus C1, adenovirus Ad40 and adenovirus SA18.

47. The method according to claim 1, wherein the first adenovirus is Pan5.

\* \* \* \* \*